

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 29, 2004, 14:53:14 ; Search time 331.622 Seconds
(without alignments)
9184.983 Million cell updates/sec

Title: US-09-989-981A-9_COPY_3_104
Perfect score: 102
Sequence: 1 ctggtaggtgagatctctga.....aacaagctgtcctggaggcc 102

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*

28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		%					ID	Description
Result	Query	Match	Length	DB				
No.	Score							
c 1	102	100.0	435	9	AI574075		AI574075	uj67h11.y
c 2	102	100.0	500	9	AI151811		AI151811	ui46c10.y
c 3	102	100.0	510	10	BB610072		BB610072	BB610072
c 4	102	100.0	511	9	AI157365		AI157365	ui45h01.y
c 5	102	100.0	583	13	BY705076		BY705076	BY705076
c 6	102	100.0	2417	11	AK050938		AK050938	Mus muscu
c 7	102	100.0	3623	11	AK004871		AK004871	Mus muscu
c 8	92.4	90.6	303	10	BB870338		BB870338	BB870338
c 9	48.6	47.6	1003	29	CNS02S18		AL211301	Tetraodon
c 10	42.6	41.8	393	10	BB870541		BB870541	BB870541
c 11	39.8	39.0	306	10	BB605863		BB605863	BB605863
c 12	33.2	32.5	522	12	BM735433		BM735433	MONO1_20_
c 13	32.2	31.6	597	9	AA524439		AA524439	ng44f07.s
c 14	32.2	31.6	752	13	BX506811		BX506811	DKFZp779M
c 15	32.2	31.6	4443	29	AY399795		AY399795	Homo sapi
c 16	31.2	30.6	500	10	BF483989		BF483989	WHE2306_H
c 17	31.2	30.6	693	10	BE404165		BE404165	WHE1201_H
c 18	31	30.4	648	28	BH295020		BH295020	CH230-44L
c 19	30.4	29.8	238	9	AV277244		AV277244	AV277244
c 20	30.4	29.8	294	9	AA706660		AA706660	ag90h11.r
c 21	30.4	29.8	295	12	BG980021		BG980021	CM3-CN009
c 22	30.4	29.8	986	9	AV254401		AV254401	AV254401
c 23	30	29.4	746	29	CC921947		CC921947	t060j23ba
c 24	29.8	29.2	551	9	AU180833		AU180833	AU180833
c 25	29.6	29.0	444	10	BF473385		BF473385	WHE0923_H
c 26	29.6	29.0	483	13	BQ467131		BQ467131	HS02L11r
c 27	29.6	29.0	518	14	CD912921		CD912921	G550.116E
c 28	29.6	29.0	564	12	BG606129		BG606129	WHE2960_H
c 29	29.6	29.0	661	12	BM377546		BM377546	EBem04_SQ
c 30	29.6	29.0	697	13	BQ466828		BQ466828	HS01L11T
c 31	29.6	29.0	730	13	BQ838111		BQ838111	WHE2906_F
c 32	29.4	28.8	429	14	CB360743		CB360743	ZF001-P00
c 33	29.4	28.8	463	13	BQ993297		BQ993297	QGF28E04.
c 34	29.4	28.8	1021	29	CNS02AAN		AL188312	Tetraodon
c 35	29.2	28.6	536	14	CA628204		CA628204	wle1.pk00
c 36	29.2	28.6	674	13	BQ743419		BQ743419	WHE4103_G
c 37	29	28.4	362	28	BH489764		BH489764	BOHQC87TR
c 38	29	28.4	436	13	BU046816		BU046816	PP_LEa002
c 39	29	28.4	614	13	BU042469		BU042469	PP_LEa001
c 40	29	28.4	630	13	BU046581		BU046581	PP_LEa002
c 41	29	28.4	635	13	BU044321		BU044321	PP_LEa001
c 42	29	28.4	735	28	BH109216		BH109216	RPCI-24-3
c 43	28.8	28.2	342	9	AI117880		AI117880	uc41f02.r
c 44	28.8	28.2	398	9	AA177634		AA177634	mt32h12.r
c 45	28.8	28.2	416	12	BG550348		BG550348	947039G04

	46	28.8	28.2	510	13	BQ557757	BQ557757 H4048B01-
c	47	28.8	28.2	524	13	BX514645	BX514645 BX514645
c	48	28.8	28.2	536	13	BX520764	BX520764 BX520764
c	49	28.8	28.2	598	9	AI591944	AI591944 mt32h12.y
c	50	28.8	28.2	654	29	DR36H15S	AL987137 Danio rer

ALIGNMENTS

RESULT 1

AI574075/c

LOCUS AI574075 435 bp mRNA linear EST 29-MAR-1999

DEFINITION uj67h11.y1 Sugano mouse liver mlia Mus musculus cDNA clone

IMAGE:1925061 5', mRNA sequence.

ACCESSION AI574075

VERSION AI574075.1 GI:4537449

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 435)

AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

TITLE The WashU-NCI Mouse EST Project 1999

JOURNAL Unpublished (1999)

COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:981353

Seq primer: custom primer used

High quality sequence stop: 432.

FEATURES

source

Location/Qualifiers

1. .435

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL"

/db_xref="taxon:10090"

/clone="IMAGE:1925061"

/sex="female"

/dev_stage="adult"

/lab_host="DH10B"

/clone_lib="Sugano mouse liver mlia"

/note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer

[ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was

ligated to a DraIII adaptor [TGTTGGCCTACTGG], digested

and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end primer CGACCTGCAGCTCGAGCACA."

ORIGIN

Query Match 100.0%; Score 102; DB 9; Length 435;
 Best Local Similarity 100.0%; Pred. No. 3.1e-22;
 Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 166 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 107
 Qy 61 ACTGTTGTCACTTTCCGAGGAGAACAAAGCTGTCCTGGAGGCC 102
 ||||||||||||||||||||||||||||||||||||||||
 Db 106 ACTGTTGTCACTTTCCGAGGAGAACAAAGCTGTCCTGGAGGCC 65

RESULT 2

AI151811/c

LOCUS AI151811 500 bp mRNA linear EST 30-SEP-1998

DEFINITION ui46cl0.y1 Sugano mouse embryo mewa Mus musculus cDNA clone
 IMAGE:1885458 5', mRNA sequence.

ACCESSION AI151811

VERSION AI151811.1 GI:3680280

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 500)

AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.

TITLE The WashU-HHMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of MedicineP

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:969782

Seq primer: custom primer used

High quality sequence stop: 499.

FEATURES Location/Qualifiers

source 1. .500


```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1885458"
/dev_stage="embryo, 14 dpc"
/lab_host="DH10B"
/clone_lib="Sugano mouse embryo mewa"
/note="Vector: pME18S-FL3; Site_1: DraIII (CACTGTGTG);
Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed
with an oligo(dT) primer [ATGTGGCCTTTTTTTTTTTTTTTTTT];
double-stranded cDNA was ligated to a DraIII adaptor
[TGTTGGCCTACTGG], digested and cloned into distinct DraIII
sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site
CACCATGTG). XhoI should be used to isolate the cDNA
insert. Size selection was performed to exclude fragments
<1.5kb. Library constructed by Dr. Sumio Sugano
(University of Tokyo Institute of Medical Science).
Custom primers for sequencing: 5' end primer
CTTCTGCTCTAAAAGCTGCG and 3' end primer
CGACCTGCAGCTCGAGCACA."

```

ORIGIN

```

Query Match          100.0%;  Score 102;  DB 9;  Length 500;
Best Local Similarity 100.0%;  Pred. No. 3.4e-22;
Matches 102;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
        |||
Db      228 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 169

Qy      61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 102
        |||
Db      168 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 127

```

RESULT 3

BB610072/c

LOCUS BB610072 510 bp mRNA linear EST 26-OCT-2001
DEFINITION BB610072 RIKEN full-length enriched, adult male liver Mus musculus
cDNA clone 1300007N20 5', mRNA sequence.

ACCESSION BB610072

VERSION BB610072.1 GI:16451685

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 510)

AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

JOURNAL Unpublished (2001)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. . 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,
Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.

FEATURES Location/Qualifiers
source 1. .510
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1300007N20"
/sex="male"
/tissue_type="liver"
/dev_stage="adult"
/clone_lib="RIKEN full-length enriched, adult male liver"

ORIGIN

Query Match 100.0%; Score 102; DB 10; Length 510;
Best Local Similarity 100.0%; Pred. No. 3.4e-22;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
|||||
Db 228 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 169
Qy 61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 102

|||||
Db 168 ACTGTTGTCACTTTCCGAGGAGACAAGCTGTCCTGGAGGCC 127

RESULT 4

AI157365/c

LOCUS AI157365 511 bp mRNA linear EST 30-SEP-1998

DEFINITION ui45h01.y1 Sugano mouse embryo mewa Mus musculus cDNA clone
IMAGE:1885393 5', mRNA sequence.

ACCESSION AI157365

VERSION AI157365.1 GI:3685834

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 511)

AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The WashU-HHMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of MedicineP

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:969717

Seq primer: custom primer used

High quality sequence stop: 480.

FEATURES

source

Location/Qualifiers

1. .511

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL"

/db_xref="taxon:10090"

/clone="IMAGE:1885393"

/dev_stage="embryo, 14 dpc"

/lab_host="DH10B"

/clone_lib="Sugano mouse embryo mewa"

/note="Vector: pME18S-FL3; Site_1: DraIII (CACTGTGTG);

Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed

with an oligo(dT) primer [ATGTGGCCTTTTTTTTTTTTTTTT];

double-stranded cDNA was ligated to a DraIII adaptor

[TGTGGCCTACTGG], digested and cloned into distinct DraIII

sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site

CACCATGTG). XhoI should be used to isolate the cDNA

insert. Size selection was performed to exclude fragments

<1.5kb. Library constructed by Dr. Sumio Sugano

(University of Tokyo Institute of Medical Science).

Custom primers for sequencing: 5' end primer

CTTCTGCTCTAAAAGCTGCG and 3' end primer
CGACCTGCAGCTCGAGCACA."

ORIGIN

Query Match 100.0%; Score 102; DB 9; Length 511;
Best Local Similarity 100.0%; Pred. No. 3.4e-22;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
|||||
Db 221 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 162

Qy 61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 102
|||||
Db 161 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 120

RESULT 5

BY705076/c

LOCUS BY705076 583 bp mRNA linear EST 16-DEC-2002

DEFINITION BY705076 RIKEN full-length enriched, adult male liver Mus musculus
cDNA clone 1300003C16 5', mRNA sequence.

ACCESSION BY705076

VERSION BY705076.1 GI:27116215

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 583)

AUTHORS Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
Kawaji,H., Kawasaki,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
Rogers,J., Birney,E. and Hayashizaki,Y.

TITLE Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 MEDLINE 22354683
 PUBMED 12466851
 COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp,
 URL:http://genome.gsc.riken.go.jp/
 Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
 Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
 Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
 Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
 Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
 Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
 Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
 Direct Submission
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.
 FEATURES Location/Qualifiers
 source 1. .583
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="1300003C16"
 /sex="male"
 /tissue_type="liver"
 /dev_stage="adult"
 /clone_lib="RIKEN full-length enriched, adult male liver"
 ORIGIN
 Query Match 100.0%; Score 102; DB 13; Length 583;
 Best Local Similarity 100.0%; Pred. No. 3.7e-22;
 Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAG 60
 |||
 Db 236 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAG 177

Qy 61 ACTGTTGTCACTTTCCGAGGAGACAAGCTGTCCTGGAGGCC 102
 |||
 Db 176 ACTGTTGTCACTTTCCGAGGAGACAAGCTGTCCTGGAGGCC 135

RESULT 6

AK050938/c

LOCUS AK050938 2417 bp mRNA linear HTC 20-SEP-2003

DEFINITION Mus musculus 9 days embryo whole body cDNA, RIKEN full-length enriched library, clone:D030040P06 product:ATP-BINDING CASSETTE, SUB-FAMILY G, MEMBER 8 (STEROLIN-2) homolog [Mus musculus], full insert sequence.

ACCESSION AK050938

VERSION AK050938.1 GI:26094211

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1
 AUTHORS Carninci,P. and Hayashizaki,Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE

2
 AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159

REFERENCE

3
 AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861

REFERENCE

4
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)

REFERENCE

5
 AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 2417)

AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:http://genome.gsc.riken.go.jp/ URL:http://fantom.gsc.riken.go.jp/.

FEATURES Location/Qualifiers

source 1. .2417

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:D030040P06"

/db_xref="MGI:2418860"

/db_xref="taxon:10090"

/clone="D030040P06"

/tissue_type="whole body"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="9 days embryo"

misc_feature 1. .2417

/note="ATP-BINDING CASSETTE, SUB-FAMILY G, MEMBER 8 (STEROLIN-2) homolog [Mus musculus] (SWISSPROT|Q9DBM0, evidence: FASTY, 92%ID, 96.7%length, match=1796)"

ORIGIN

Query Match 100.0%; Score 102; DB 11; Length 2417;

Best Local Similarity 100.0%; Pred. No. 7.7e-22;

Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60

|||||

Db 186 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGGTGGACTGACCACTGTAGGTGAAGTACAG 127

Qy 61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 102
 |||||

Db 126 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 85

RESULT 7

AK004871/c

LOCUS AK004871 3623 bp mRNA linear HTC 20-SEP-2003

DEFINITION Mus musculus adult male liver cDNA, RIKEN full-length enriched library, clone:1300003C16 product:ATP-BINDING CASSETTE, SUB-FAMILY G, MEMBER 8 (STEROLIN-2) homolog [Mus musculus], full insert sequence.

ACCESSION AK004871

VERSION AK004871.1 GI:12836380

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1
 AUTHORS Carninci,P. and Hayashizaki,Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE

2
 AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159

REFERENCE

3
 AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861

REFERENCE

4
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)

REFERENCE

5
 AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 3623)

AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGCGCCGCAACTCGAGTTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence[5' GAGAGAGAGAAGGATCCAAGAGCTCAATTAATTTAATTAAACCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end: XhoI. Host: SOLR.

FEATURES

Location/Qualifiers

source 1. .3623

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:1300003C16"

/db_xref="MGI:1896857"

/db_xref="taxon:10090"

/clone="1300003C16"

/sex="male"

/tissue_type="liver"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

CDS 69. .2090

/note="unnamed protein product; ATP-BINDING CASSETTE, SUB-FAMILY G, MEMBER 8 (STEROLIN-2) homolog [Mus musculus] (SWISSPROT|Q9DBM0, evidence: FASTY, 92%ID, 96.7%length, match=1796)

putative"

/codon_start=1

```

/protein_id="BAB23630.1"
/db_xref="GI:12836381"
/translation="MAEKTKEETQLWNGTVLQDASQGLQDSLFSSSEDNSLYFTYSGQ
SNTLEVRDLTYQVDIASQVPWFEEQLAQFKIPWRSHSSQDSCELGIRNLSFKVRSQML
AIIGSSGCGRASLLDVITGRGHGGKMKSGQIWINGQPSTPQLVRKCVAHVRQHDQLLP
NLTVRETLAFIAQMRLPRTFSQAQRDKRVEDVIAELRLRQCANTRVGNTYVRGVSGGE
RRRVSIGVQLLWNPGLILDEPTSGLDSTAHNLVTTLSRLAKGNRLVLISLHQPRSD
IFRLEFDLVLLMTSGTPIYLGAAQOMVQYFTSIGHPCPRYSNPADFYVDLTSIDRRSKE
REVATVEKAQSLAALFLEKVQGFDDFLWKAEAKELNTSTHTVSLTLTQDTCGTAVEL
PGMIEQFSTLIRROIENDFRDLPTLLIHGSEACLSLIIGFLYYGHGAKQLSFMDTAA
LLFMIGALIPFNVILDVVSCKHSERSMLYYELEDGLYTAGPYFFAKILGELPEHCAYV
IIYAMPIYWLTLNLRPVPPELFLHFLLVVLFCCRTMALAASAMLPFHMSSFFCNAL
YNSFYLTAGFMINLDNLWIVPAWISKLSFLRWCFSGLMQIQFNHLYTTQIGNFTFSI
LGDTMISAMDNLNSHPLYAIYLIVIGISYGFLLFLYYLSLKLKQKSIQDW"

polyA_signal 3605. .3610
               /note="putative"
polyA_site    3623
               /note="putative"

```

ORIGIN

```

Query Match          100.0%;   Score 102;   DB 11;   Length 3623;
Best Local Similarity 100.0%;   Pred. No. 9.5e-22;
Matches 102;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy      1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      236 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 177

Qy      61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 102
        ||||||||||||||||||||||||||||||||||||||||
Db      176 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 135

```

RESULT 8

BB870338/c

LOCUS BB870338 303 bp mRNA linear EST 27-NOV-2001

DEFINITION BB870338 RIKEN full-length enriched, adult male intestinal mucosa
Mus musculus cDNA clone G630020H06 5', mRNA sequence.

ACCESSION BB870338

VERSION BB870338.1 GI:17116548

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 303)

AUTHORS Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,
Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,
Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.

TITLE RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)

JOURNAL Unpublished (2001)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. . 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
e mouse tissues.

FEATURES Location/Qualifiers
source 1. .303
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G630020H06"
/sex="male"
/tissue_type="intestinal mucosa"
/dev_stage="adult"
/clone_lib="RIKEN full-length enriched, adult male
intestinal mucosa"

ORIGIN

Query Match 90.6%; Score 92.4; DB 10; Length 303;
Best Local Similarity 94.1%; Pred. No. 3.3e-19;
Matches 96; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
|||||
Db 281 CTGGTAGGTGAGATCTCTGACCTCCAGAGGGTGGACTGACCACTGTAGTTGAAGTACAG 222
QY 61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 102
|||||
Db 221 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTTGATGGC 180

RESULT 9
 CNS02S18
 LOCUS CNS02S18 1003 bp DNA linear GSS 01-SEP-2000
 DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone
 161A20 of library G from Tetraodon nigroviridis, genomic survey
 sequence.
 ACCESSION AL211301
 VERSION AL211301.1 GI:7870120
 KEYWORDS GSS; genome survey sequence.
 SOURCE Tetraodon nigroviridis
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodontidae; Tetraodon.
 REFERENCE 1
 AUTHORS Roest Crollius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,
 Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,
 Saurin, W. and Weissenbach, J.
 TITLE Estimate of human gene number provided by genome-wide analysis
 using Tetraodon nigroviridis DNA sequence
 JOURNAL Nat. Genet. 25 (2), 235-238 (2000)
 MEDLINE 20296633
 PUBMED 10835645
 REFERENCE 2
 AUTHORS Roest Crollius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C.,
 Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,
 Saurin, W., Bernot, A. and Weissenbach, J.
 TITLE Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 JOURNAL Genome Res. 10 (7), 939-949 (2000)
 MEDLINE 20359837
 PUBMED 10899143
 REFERENCE 3 (bases 1 to 1003)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetraodon>.
 FEATURES Location/Qualifiers
 source 1. .1003
 /organism="Tetraodon nigroviridis"
 /mol_type="genomic DNA"
 /db_xref="taxon:99883"
 /clone="161A20"
 /clone_lib="G"
 /note="Genoscope sequence ID : C0AG161BA10LP1~end : T7"
 ORIGIN
 Query Match 47.6%; Score 48.6; DB 29; Length 1003;
 Best Local Similarity 68.0%; Pred. No. 8.5e-05;
 Matches 66; Conservative 1; Mismatches 30; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
 || ||| |||| || ||||| ||| || || ||||| ||||| ||
 Db 579 CTCATAGTTGAGGTCGTTGACCTCCAGCTCGTTGCACCCTCCGCTGTAGGTGAAGTAGAG 638

Qy 61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGG 97
 ||| ||| ||||| ||| :||| ||
 Db 639 GCTGCTGTWCTCCTCCGAGAACAGCTGTMTGTCTTTG 675

RESULT 10
 BB870541/c

LOCUS BB870541 393 bp mRNA linear EST 27-NOV-2001
 DEFINITION BB870541 RIKEN full-length enriched, adult male intestinal mucosa
 Mus musculus cDNA clone G630022C22 5', mRNA sequence.

ACCESSION BB870541

VERSION BB870541.1 GI:17116751

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 393)

AUTHORS Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
 Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,
 Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
 Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,
 Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,
 Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
 Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
 Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A.,
 Muramatsu,M. and Hayashizaki,Y.

TITLE RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
 2001)

JOURNAL Unpublished (2001)

COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp,
 URL:http://genome.gsc.riken.go.jp/
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. . 10 (10), 1617-1630 (2000)
 wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
 Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
 Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
 and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
 10 (11), 1757-1771 (2000)
 Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
 Sugahara,Y. and Hayashizaki,Y.

Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
e mouse tissues.

FEATURES Location/Qualifiers
 source 1. .393
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="G630022C22"
 /sex="male"
 /tissue_type="intestinal mucosa"
 /dev_stage="adult"
 /clone_lib="RIKEN full-length enriched, adult male
 intestinal mucosa"

ORIGIN

Query Match 41.8%; Score 42.6; DB 10; Length 393;
Best Local Similarity 76.2%; Pred. No. 0.0045;
Matches 80; Conservative 0; Mismatches 19; Indels 6; Gaps 2;

Qy 4 GTAGGTGAGATCTCTGACCTCCAGAGTG---TTGGACTGACCACTGTAGGTGAAGTACAG 60
 ||| || ||||| ||||| |||| | | | | ||||| |||||
Db 318 GTATGTTAGATCTCTTACCTCCATAGTGTTTGTGAGCTTACCAGCTCTAGGTTAAGTACAG 259

Qy 61 ACTGTTGTCACTTTCCGAGGAGAAC---AAGCTGTCCTGGAGGCC 102
 ||||| ||||| ||||| | | ||||| ||||| |
Db 258 ACTGTTGTCACTTTCCCTAGGAGTAAGCAAGGCTGTCCTGGAGGGC 214

RESULT 11

BB605863/c

LOCUS BB605863 306 bp mRNA linear EST 05-DEC-2000

DEFINITION BB605863 RIKEN full-length enriched, 0 day neonate lung Mus
 musculus cDNA clone E030013I04 5', mRNA sequence.

ACCESSION BB605863

VERSION BB605863.1 GI:11557265

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

 ORGANISM Mus musculus

 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 306)

 AUTHORS Aizawa,K., Akahira,S., Akimura,T., Arai,A., Arakawa,T.,
 Carninci,P., Hanagaki,T., Hayatsu,N., Hiraoka,T., Hirozane,T.,
 Hodoyama,Y., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kawai,J.,
 Kojima,Y., Konno,H., Kusakabe,M., Matsuyama,T., Miyazaki,A.,
 Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Okazaki,Y.,
 Okido,T., Owa,C., Sakai,C., Sakai,K., Sasaki,D., Sato,K.,
 Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
 Suzuki,H., Tagawa,A., Takahashi,F., Tanaka,T., Toya,T.,
 Watahiki,A., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshiki,A.,
 Muramatsu,M. and Hayashizaki,Y.

 TITLE RIKEN Mouse ESTs (Aizawa,K. et al. 2000)

JOURNAL Unpublished (2000)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S.,
Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermolabile enzymes by
trehalose and its application for the synthesis of full length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,
Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for
further details.

FEATURES Location/Qualifiers
source 1. .306
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="E030013I04"
/tissue_type="lung"
/dev_stage="0 day neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 0 day neonate
lung"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGCGGCCGCAACTCGAGTTTTTTTTTTTTTTTTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Query Match 39.0%; Score 39.8; DB 10; Length 306;
Best Local Similarity 74.6%; Pred. No. 0.032;
Matches 50; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 36 ACTGACCACTGTAGGTGAAGTACAGACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCT 95
 || ||| | |||| || ||| | ||| ||||| ||||| || | || |||
 Db 304 ACCCACCTCCGTAGTTGCAGTTCCAACCTCTTGTC AATTTCCGAGGAGCACCACCTATCCA 245
 Qy 96 GGAGGCC 102
 |||| ||
 Db 244 GGAGCCC 238

RESULT 12

BM735433

LOCUS BM735433 522 bp mRNA linear EST 01-MAR-2002

DEFINITION MONO1_20_F01.g1_A005 Monocytes (MONO1) Equus caballus cDNA, mRNA sequence.

ACCESSION BM735433

VERSION BM735433.1 GI:19056766

KEYWORDS EST.

SOURCE Equus caballus (horse)

ORGANISM Equus caballus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

REFERENCE 1 (bases 1 to 522)

AUTHORS Vandenplas,M.L., Cordonnier-Pratt,M.-M., Sudman,M.L., Wentzel,V.E., Gingle,A.R., Pratt,L.H. and Moore,J.N.

TITLE An EST database from equine (Equus caballus) monocytes

JOURNAL Unpublished (2001)

COMMENT Contact: Cordonnier-Pratt MM

Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 583 0210

Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for high quality sequence is 20. Three-prime sequences, which are obtained with PolyTMix or T7 sequencing primer, are presented as the reverse complement.

Seq primer: T7

High quality sequence start: 43

High quality sequence stop: 522

POLYA=Yes.

FEATURES Location/Qualifiers

source

1. .522

/organism="Equus caballus"

/mol_type="mRNA"

/db_xref="taxon:9796"

/cell_type="Isolated peripheral blood monocytes stimulated with E. coli lipopolysaccharide"

/clone_lib="Monocytes (MONO1)"

/note="Vector: pBluescript SK(-) from Lambda ZapII;

Site_1: XhoI; Site_2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda ZAPII. Clones to be sequenced were prepared by mass excision."

ORIGIN

Query Match 32.5%; Score 33.2; DB 12; Length 522;

Best Local Similarity 67.1%; Pred. No. 5.7;


```

Matches    47;  Conservative      0;  Mismatches    23;  Indels      0;  Gaps      0;

Qy          6 AGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTGT 65
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db          200 AGTTCAAATATTGCCTGTCCAGAGAGGTTGTCCGACCACTGTAGCTGAAGCAGCGTCTCC 259

Qy          66 TGTCACTTTC 75
           ||| ||| ||| ||| ||| |||
Db          260 AGTCACTTTC 269

```

I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. "

ORIGIN

Query Match 31.6%; Score 32.2; DB 9; Length 597;
Best Local Similarity 63.6%; Pred. No. 13;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
||| ||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| |||
Db 112 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTTCATTCTG 53

Qy 65 TTGTCACCTTTCCGAGGA 81
|| ||| ||| | |||
Db 52 TTCTCAGTTTTCTGGA 36

RESULT 14

BX506811

LOCUS BX506811 752 bp mRNA linear EST 04-SEP-2003

DEFINITION DKFZp779M191_r1 779 (synonym: hnccl) Homo sapiens cDNA clone
DKFZp779M191 5', mRNA sequence.

ACCESSION BX506811

VERSION BX506811.1 GI:32047420

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 752)

AUTHORS Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.

TITLE EST (Poustka,A., Albert,R., Moosmayer,P., Schupp,I.,
Wellenreuther,R., et al.)

JOURNAL Unpublished (2003)

COMMENT Contact: MIPS

MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.

No sl sequence available.

This clone (DKFZp779M191) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES Location/Qualifiers

source 1. .752

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="DKFZp779M191"

/tissue_type="liver"

```

/dev_stage="fetal"
/lab_host="DH10B"
/clone_lib="779 (synonym: hnccl)"
/note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB"

```

ORIGIN

```

Query Match          31.6%;  Score 32.2;  DB 13;  Length 752;
Best Local Similarity 63.6%;  Pred. No. 15;
Matches   49;  Conservative   0;  Mismatches   28;  Indels     0;  Gaps     0;

```

```

Qy          5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
             || | |||| ||      ||| ||||| ||  | | ||| || | ||| | || |||
Db          303 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTCATTCTG 362

Qy          65 TTGTCACTTTCCGAGGA 81
             || ||| ||| | |||
Db          363 TTCTCAGTTTTCCTGGA 379

```

RESULT 15

```

AY399795
LOCUS      AY399795                4443 bp    DNA        linear    GSS 15-DEC-2003
DEFINITION Homo sapiens CFTR gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION  AY399795
VERSION    AY399795.1  GI:39755784
KEYWORDS   GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 4443)
AUTHORS     Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
TITLE       Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
JOURNAL      Science 302 (5652), 1960-1963 (2003)
PUBMED      14671302
REFERENCE   2 (bases 1 to 4443)
AUTHORS     Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
TITLE       Direct Submission
JOURNAL      Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source            1..4443
                       /organism="Homo sapiens"
                       /mol_type="genomic DNA"
                       /db_xref="taxon:9606"
     gene              <1..>4443
                       /gene="CFTR"

```

/locus_tag="HCM0343"

ORIGIN

Query Match 31.6%; Score 32.2; DB 29; Length 4443;
Best Local Similarity 63.6%; Pred. No. 37;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
||| ||||| || ||| ||||| || | | | ||| || | ||| | || |||
Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTTTCATTCTG 1472

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 1473 TTCTCAGTTTTCCTGGA 1489

RESULT 16

BF483989/c

LOCUS BF483989 500 bp mRNA linear EST 06-DEC-2000

DEFINITION WHE2306_H10_O20ZS Wheat pre-anthesis spike cDNA library *Triticum aestivum* cDNA clone WHE2306_H10_O20, mRNA sequence.

ACCESSION BF483989

VERSION BF483989.1 GI:11567278

KEYWORDS EST.

SOURCE *Triticum aestivum* (bread wheat)

ORGANISM *Triticum aestivum*

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; *Triticum*.

REFERENCE 1 (bases 1 to 500)

AUTHORS Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D.,
Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J.,
Seaton, C.L. and Tong, J.C.

TITLE The structure and function of the expressed portion of the wheat
genomes - Pre-anthesis spike cDNA library

JOURNAL Unpublished (2000)

COMMENT Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105595773

Fax: 5105595818

Email: oandersn@pw.usda.gov

Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20

Seq primer: Stratagene SK primer.

FEATURES

source

Location/Qualifiers

1..500

/organism="Triticum aestivum"

/mol_type="mRNA"

/cultivar="Chinese Spring"

/db_xref="taxon:4565"

/clone="WHE2306_H10_O20"

/tissue_type="Spike before anthesis"

/dev_stage="Adult plant"

/lab_host="E. coli SOLR"

/clone_lib="Wheat pre-anthesis spike cDNA library"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
greenhouse. Whole spike with awns trimmed, white, green
and yellow anther were collected and total RNA, and
poly(A) RNA were prepared, a cDNA library was made, and
the cDNA clones were in vivo excised to give pBluescript
phagemids in the TJ Close lab (Choi, Close, Fenton) at
the University of California, Riverside. Plasmid DNA
preparations and DNA sequencing were performed in the OD
Anderson lab (all other authors)."

ORIGIN

Query Match 30.6%; Score 31.2; DB 10; Length 500;
Best Local Similarity 57.0%; Pred. No. 25;
Matches 57; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
||||| | ||| ||||| | ||| | ||| | ||| ||| ||| |||
Db 389 CTGGTGCCCGCAATCCTGCACCTCGATGGTGCACGCCTGGTCGAAGCAGATGCAGCACAG 330

Qy 61 ACTGTTGTCACTTTCCGAGGAGAACAAAGCTGTCCTGGAGG 100
|||| || || ||||| || ||| ||| |
Db 329 CTCCGTGTCGCTCACCTCCGAGAACATGTCGTCGTCGATG 290

RESULT 17

BE404165/c

LOCUS BE404165 693 bp mRNA linear EST 21-JUL-2000

DEFINITION WHE1201_H12_O23ZS Wheat etiolated seedling root cDNA library
Triticum aestivum cDNA clone WHE1201_H12_O23, mRNA sequence.

ACCESSION BE404165

VERSION BE404165.1 GI:9363633

KEYWORDS EST.

SOURCE Triticum aestivum (bread wheat)

ORGANISM Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.

REFERENCE 1 (bases 1 to 693)

AUTHORS Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D.,
Han,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
Seaton,C.L. and Tong,J.C.

TITLE The structure and function of the expressed portion of the wheat
genomes

JOURNAL Unpublished (2000)

COMMENT Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center

800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105595773

Fax: 5105595818

Email: oandersn@pw.usda.gov

Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20

Seq primer: Strategene SK primer.

FEATURES

Location/Qualifiers

source 1..693

```

/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE1201_H12_O23"
/tissue_type="Root"
/dev_stage="Five day old etiolated seedling"
/lab_host="E. coli SOLR"
/clone_lib="Wheat etiolated seedling root cDNA library"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Seeds were
surface-sterilized, germinated and grown aseptically in
the dark at room temperature on filter paper with water,
nystatin and cefotaxime in covered crystallization
dishes. Roots were harvested. The tissue, total RNA, and
poly(A) RNA were prepared, a cDNA library was made, and
the cDNA clones were in vivo excised to give pBluescript
phagemids in the TJ Close lab (Choi, Close, Fenton) at the
University of California, Riverside. Plasmid DNA
preparations and DNA sequencing were performed in the OD
Anderson lab (all other authors)."
```

ORIGIN

```

Query Match          30.6%; Score 31.2; DB 10; Length 693;
Best Local Similarity 57.0%; Pred. No. 29;
Matches 57; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy      1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
      |||||  | |||  ||||| | |||  | |||  | ||| |||||
Db      319 CTGGTGCCCGCAATCCTGCACCTCGATGGTGCACGCCTGGTCGAAGCAGATGCAGCACAG 260

Qy      61 ACTGTTGTCACTTTCCGAGGAGAACAAAGCTGTCCTGGAGG 100
      |||| || ||  ||||| | ||| | || |
Db      259 CTCCGTGTCGCTCACCTCCGAGAACATGTCGTCGTCGATG 220
```

RESULT 18

BH295020/c

LOCUS BH295020 648 bp DNA linear GSS 30-NOV-2001

DEFINITION CH230-44L24.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone
CH230-44L24, genomic survey sequence.

ACCESSION BH295020

VERSION BH295020.1 GI:17207428

KEYWORDS GSS.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 648)

AUTHORS Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.

TITLE Rat BAC End Sequences from Library CHORI-230 EcoRI segment

JOURNAL Unpublished (1999)

COMMENT Other_GSSs: CH230-44L24.TV
Contact: Shaying Zhao

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(<http://www.chori.org/bacpac/rat230.htm>). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/orering_information.htm). BAC end
page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 44 row: L column: 24
Seq primer: SP6
Class: BAC ends.

FEATURES
source Location/Qualifiers
1. .648
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SsNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-44L24"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 1"
/note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
Pieter de Jong"

ORIGIN

Query Match 30.4%; Score 31; DB 28; Length 648;
Best Local Similarity 59.8%; Pred. No. 33;
Matches 52; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 8 GTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTGTTG 67
||||| ||| | |||| || | ||| | || | || ||||| | |
Db 355 GTGAGTTCTGTACCTGCAAGGCCAAAGGCCAAAGCAGAAAACTAAACTACAGAGGGAAG 296
Qy 68 TCACTTTCCGAGGAGAACAAGCTGTCC 94
|| ||||| ||||| || ||
Db 295 TCTCTTTCTTGTGAGAACATTCTCACC 269

RESULT 19

AV277244/c

LOCUS AV277244 238 bp mRNA linear EST 05-NOV-1999

DEFINITION AV277244 RIKEN full-length enriched, adult male testis (DH10B) Mus
musculus cDNA clone 4932441F23 3', mRNA sequence.

ACCESSION AV277244

VERSION AV277244.1 GI:6265281

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 238)

AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T.,

TITLE
JOURNAL
COMMENT

Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F.,
Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I.,
Kai,C., Kawai,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M.,
Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K.,
Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y.,
Suzuki,H., Suzuki,H., Takahashi,F., Tateno,M., Tominaga,N.,
Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yasunishi,A.,
Yokota,T., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Konno,H., et al. 1999)

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:<http://genome.gsc.riken.go.jp/>
Sasaki,N., Izawa,M., Watahiki,M., Ozawa,K., Tanaka,T., Yoneda,Y.,
Matsuura,S., Carninci,P., Muramatsu,M., Okazaki,Y. and
Hayashizaki,Y.

Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)

Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,
Okazaki,Y. and Hayashizaki,Y.

Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)

Carninci,P. and Hayashizaki,Y.

High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)

Please visit our web site (<http://genome.rtc.riken.go.jp>) for
further details.

FEATURES
source

Location/Qualifiers

1. .238
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="4932441F23"
/sex="male"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult male testis
(DH10B)"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse

transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI."

ORIGIN

Query Match 29.8%; Score 30.4; DB 9; Length 238;
 Best Local Similarity 59.1%; Pred. No. 30;
 Matches 52; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 14 TCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTGTTGTCACCTT 73
 | ||| | || ||| | | ||| ||| | | |||| | || ||||
 Db 222 TTTCTCAACTATAGAATCTAGTTGTGAAGACTTTTCATAAAGTTGCTCTTGAAAACACTT 163

Qy 74 TCCGAGGAGACAAGCTGTCCTGGAGGC 101
 | ||| ||||| |||| || | ||
 Db 162 TTCGATAAGAACAATCTGTTCTGTAGC 135

RESULT 20

AA706660

LOCUS AA706660 294 bp mRNA linear EST 24-DEC-1997

DEFINITION ag90h11.rl Stratagene hNT neuron (#937233) Homo sapiens cDNA clone IMAGE:1141797 5', mRNA sequence.

ACCESSION AA706660

VERSION AA706660.1 GI:2716578

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 294)

AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.

TITLE WashU-NCI human EST Project

JOURNAL Unpublished (1997)

COMMENT Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -28m13 rev1 ET from Amersham

High quality sequence stop: 281.

FEATURES

source

Location/Qualifiers

1. .294

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:1141797"

```

/dev_stage="hNT neurons"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene hNT neuron (#937233)"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dT.
Differentiated, post mitotic hNT neurons. Average insert
size: 1.5 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5'
GAATTCGGCAGAG 3' ~3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTTTTTT 3'"

```

ORIGIN

```

Query Match          29.8%;  Score 30.4;  DB 9;  Length 294;
Best Local Similarity 59.1%;  Pred. No. 34;
Matches    52;  Conservative    0;  Mismatches    36;  Indels      0;  Gaps      0;

Qy          4 GTAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAGACT 63
             | | | | | | | | | | | | | | | | | | | | | | | |
Db          52 GAAAGAGAAATCTTTGAGCTCCTGAATGTGGAACAACCTAATGGGAGGGAAGAAGAAAAA 111

Qy          64 GTTGTCACCTTCCGAGGAGAACAAGCTG 91
             | | | | | | | | | | | |
Db          112 TTGGGGGCTTTGAAAGGAGAACAGCGTG 139

```

RESULT 21

```

BG980021/c
LOCUS      BG980021                295 bp    mRNA    linear    EST 12-JUN-2001
DEFINITION CM3-CN0092-180101-644-a01 CN0092 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BG980021
VERSION    BG980021.1  GI:14382756
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 295)
  AUTHORS  Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE    Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
  PUBMED    10737800
  COMMENT   Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL

```

(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM3&t2=CM3-CN0092-180101-644-a01&t3=2001-01-18&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 36

High quality sequence stop: 294.

FEATURES
source Location/Qualifiers
1. .295
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CN0092"
/note="Organ: colon_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Query Match 29.8%; Score 30.4; DB 12; Length 295;
Best Local Similarity 63.9%; Pred. No. 34;
Matches 46; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 10 GAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTGTTGTC 69
||||| || ||| ||||| || | | | ||| || | ||| ||| ||
Db 294 GAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTTCATTCTGTTCTC 235

Qy 70 ACTTTCCGAGGA 81
| ||| | |||
Db 234 AGTTTTCCTGGA 223

RESULT 22

AV254401/c

LOCUS AV254401 986 bp mRNA linear EST 24-OCT-2001

DEFINITION AV254401 RIKEN full-length enriched, adult male testis (DH10B) Mus
musculus cDNA clone 4921509J16 3', mRNA sequence.

ACCESSION AV254401

VERSION AV254401.2 GI:16388054

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 986)

AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.

TITLE RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

JOURNAL Unpublished (2001)

COMMENT On Nov 4, 1999 this sequence version replaced gi:6241860.

Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp,
 URL: <http://genome.gsc.riken.go.jp/>

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. *Genome Res.* . 10 (10), 1617-1630 (2000)

wagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
 Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.
 and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* .
 10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,
 Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. *Genome Res.* . 11 (2), 281-289 (2001)

Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I.,
 Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
 Hayashizaki, Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences *Mamm. Genome.* 12, 673-677 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
 further details.

cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.

FEATURES

source

Location/Qualifiers

1. .986
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="4921509J16"
 /sex="male"
 /tissue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, adult male testis
 (DH10B)"
 /note="Site_1: SalI; Site_2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was

primed with a primer [5'
GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I. Cloning sites, 5' end: Sall; 3' end: BamHI."

ORIGIN

Query Match 29.8%; Score 30.4; DB 9; Length 986;
Best Local Similarity 59.1%; Pred. No. 64;
Matches 52; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 14 TCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTGTTGTCACCTT 73
| | | | | | | | | | | | | | | | | | | | | |
Db 375 TTTCTCAACTATAGAATCTAGTGTGAAGACTTTTCATTAAGTTGCTCTTGAGAACACTT 316

Qy 74 TCCGAGGAGAACAAGCTGTCCTGGAGGC 101
| | | | | | | | | | | | | | | | | |
Db 315 TCCGATGAGAGCGATCTGTTCTTGTAGC 288

RESULT 23

CC921947/c

LOCUS CC921947 746 bp DNA linear GSS 08-AUG-2003

DEFINITION t060j23ba.fl1 TAMBT Bos taurus genomic clone t060j23ba, genomic
survey sequence.

ACCESSION CC921947

VERSION CC921947.1 GI:33555987

KEYWORDS GSS.

SOURCE Bos taurus (cow)

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 746)

AUTHORS Lin, S., Najjar, F.Z., Adelson, D., Gill, C.A. and Roe, B.A.

TITLE Bovine BAC End Sequences from Library TAMBT

JOURNAL Unpublished (2003)

COMMENT Contact: Bruce A. Roe

Advanced Center for Genome Technology

University of Oklahoma Department of Chemistry and Biochemistry

620 Parrington Oval, Room 208, Norman, OK 73019, USA

Tel: 405 325 4912

Fax: 405 325 7762

Email: broe@ou.edu

Class: BAC ends

High quality sequence start: 40

High quality sequence stop: 739.

FEATURES

source

Location/Qualifiers

1. .746

/organism="Bos taurus"

/mol_type="genomic DNA"

/strain="Angus bull T A M U Shoshone Y6 11519666"

/db_xref="taxon:9913"
/clone="t060j23ba"
/sex="Male"
/cell_type="Blood"
/clone_lib="TAMBT"
/note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
HindIII; TAMBT Bovine BAC library (Male) produced by Texas
A&M University, Department of Animal Science."

ORIGIN

Query Match 29.4%; Score 30; DB 29; Length 746;
Best Local Similarity 51.4%; Pred. No. 74;
Matches 36; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 28 AGTGTGGACTGACCACTGTAGGTGAAGTACAGACTGTTGTCACTTTCCGAGGAGAACAA 87
||||| ||||| ||| ||||| | | ||||| | |||
Db 86 AGTGCTGGAGTGATCACTGTANNNNNNNNNNNNNNNNNNNNNNTNTACGAGGACACCAA 27

Qy 88 GCTGTCCTGG 97
| | | | |
Db 26 GGTGGCTTTG 17

RESULT 24

AU180833

LOCUS AU180833 551 bp mRNA linear EST 21-MAR-2001
DEFINITION AU180833 Medaka eye cDNA library (SNK01) Oryzias latipes cDNA clone
NGY14.08d, mRNA sequence.

ACCESSION AU180833

VERSION AU180833.1 GI:13429670

KEYWORDS EST.

SOURCE Oryzias latipes (Japanese medaka)

ORGANISM Oryzias latipes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.

REFERENCE 1 (bases 1 to 551)

AUTHORS Sanaka,E., Hori,H., Naruse,K., Mitani,H. and Tanaka,M.

TITLE Medaka EST analysis

JOURNAL Unpublished (2001)

COMMENT Contact: Emi Sanaka

Department of Biological Sciences

Graduate School of Science, Nagoya University

Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan

Tel: 81-52-789-2973

Fax: 81-52-789-2974

Email: sanaka@bio.nagoya-u.ac.jp

This clone was isolated from Medaka eye cDNA library (SNK01) 5'end.

FEATURES

source

Location/Qualifiers

1. .551

/organism="Oryzias latipes"

/mol_type="mRNA"

/strain="wild type"

/db_xref="taxon:8090"

/clone="NGY14.08d"

/tissue_type="eye"

/dev_stage="adult"
/clone_lib="Medaka eye cDNA library (SNK01)"
/note="Wild samples from Okayama Pref.(Southern part of Japan)"

ORIGIN

Query Match 29.2%; Score 29.8; DB 9; Length 551;
Best Local Similarity 63.0%; Pred. No. 73;
Matches 46; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
||||| ||| ||| ||| | | | ||||| ||||| ||||| | ||| ||
Db 372 CTGGCAGGGTTGATTCTAAGATAAAAAGTGTAGGACTGATTACTGTTGAAGAAGAGAAG 431

Qy 61 ACTGTTGTCACTT 73
| | | |||
Db 432 AGTCCAGGTGCTT 444

RESULT 25

BF473385/c

LOCUS BF473385 444 bp mRNA linear EST 04-DEC-2000

DEFINITION WHE0923_H02_P03ZS Wheat 5-15 DAP spike cDNA library Triticum
aestivum cDNA clone WHE0923_H02_P03, mRNA sequence.

ACCESSION BF473385

VERSION BF473385.1 GI:11542567

KEYWORDS EST.

SOURCE Triticum aestivum (bread wheat)

ORGANISM Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.

REFERENCE 1 (bases 1 to 444)

AUTHORS Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D.,
Han,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
Seaton,C.L. and Tong,J.C.

TITLE The structure and function of the expressed portion of the wheat
genomes - 5-15 DAP spike cDNA library

JOURNAL Unpublished (2000)

COMMENT Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center

800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105595773

Fax: 5105595818

Email: oandersn@pw.usda.gov

Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20

Seq primer: Stratagene SK primer.

FEATURES

source

Location/Qualifiers

1..444

/organism="Triticum aestivum"

/mol_type="mRNA"

/cultivar="Chinese Spring"

/db_xref="taxon:4565"

/clone="WHE0923_H02_P03"

/tissue_type="Spike"

```

/dev_stage="Adult plant"
/lab_host="E. coli SOLR"
/clone_lib="Wheat 5-15 DAP spike cDNA library"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
greenhouse. Spikes at 5, 10 and 15 DAP were harvested,
total RNA and poly(A) RNA were prepared, a cDNA library
was made, and the cDNA clones were in vivo excised to
give pBluescript phagemids in the TJ Close lab (Choi,
Close, Fenton) at the University of California,
Riverside. Plasmid DNA preparations and DNA sequencing
were performed in the OD Anderson lab (all other
authors)."
```

ORIGIN

```

Query Match          29.0%;  Score 29.6;  DB 10;  Length 444;
Best Local Similarity 56.0%;  Pred. No. 76;
Matches   56;  Conservative   0;  Mismatches   44;  Indels     0;  Gaps     0;
```

```

Qy      1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
        |||||  |  |||  |||||  |  |||  |  |||  |  |||  |||  |||  |||
Db      323 CTGGTGCCCGCAATCCTGCACCTCGATGGTGCACGCCTGGTGAAGCAGATGCAGCACAG 264

Qy      61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGG 100
        ||||  ||  ||  |||||  |  |||  |  |||  |
Db      263 CTCCGTGTCGCTCACCTCCCAGAACATGTCGTCGTCGATG 224
```

RESULT 26

```

BQ467131/c
LOCUS      BQ467131                      483 bp    mRNA    linear    EST 30-MAY-2002
DEFINITION HS02L11r HS Hordeum vulgare subsp. vulgare cDNA clone HS02L11
           5-PRIME, mRNA sequence.
ACCESSION  BQ467131
VERSION    BQ467131.1  GI:21274913
KEYWORDS   EST.
SOURCE     Hordeum vulgare subsp. vulgare
  ORGANISM Hordeum vulgare subsp. vulgare
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
           Pooideae; Triticeae; Hordeum.
REFERENCE  1 (bases 1 to 483)
  AUTHORS  Zhang,H., Potokina,E., Michalek,W., Weschke,W., Stein,N. and
           Graner,A.
  TITLE    Barley ESTs from germinating seeds
  JOURNAL  Unpublished (2002)
COMMENT    Contact: Stein Nils
           Molecular Markers Group, Department Genbank
           Institute of Plant Genetics and Crop Plant Research (IPK)
           Corrensstr. 3, 06466, Gatersleben, Germany
           Tel: 039482-5522
           Fax: 039482-5595
           Email: stein@ipk-gatersleben.de
           Insert Length: 483  Std Error: 0.00
           Plate: 2  row: L  column: 11
           Seq primer: M13rev.
FEATURES   Location/Qualifiers
```



```

source      1. .483
            /organism="Hordeum vulgare subsp. vulgare"
            /mol_type="mRNA"
            /cultivar="barke"
            /sub_species="vulgare"
            /db_xref="taxon:112509"
            /clone="HS02L11"
            /tissue_type="embryo + scutellum"
            /dev_stage="0-16 hours after imbibition"
            /lab_host="XL10-Gold"
            /clone_lib="HS"
            /note="Vector: pBluescript SK+; Site_1: EcoRI (5'-end of
            cDNA); Site_2: XhoI (3'-end of cDNA); Due to a cloning
            artefact caused by the kit, in most cases the EcoRI site
            is NOT present, as well as the EcoRI adapter used for
            cloning. To excise the insert, restriction sites upstream
            EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also
            due to the cloning system used Blue/white selection for
            recombinats is not 100% reliable."

```

ORIGIN

Query Match 29.0%; Score 29.6; DB 13; Length 483;
Best Local Similarity 56.0%; Pred. No. 80;
Matches 56; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAG 60
 ||||| | || ||||| | ||| | ||| | ||| |||||
 Db 375 CTGGTGCCCGCAGTCTCTCCACCTCGATGGTGCACGCCTGGTCAAGCAGATGCAGCACAG 316

Qy 61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGG 100
 |||| | || ||||| | ||| |
 Db 315 CTCCGTGTCGCTCACCTCCGAGAACATGTCGTCGTCGACG 276

RESULT 27

CD912921/c
LOCUS CD912921 518 bp mRNA linear EST 14-JUL-2003
DEFINITION G550.116E20F010525 G550 *Triticum aestivum* cDNA clone G550116E20,
mRNA sequence.
ACCESSION CD912921
VERSION CD912921.1 GI:32687245
KEYWORDS EST.
SOURCE *Triticum aestivum* (bread wheat)
ORGANISM *Triticum aestivum*
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; *Triticum*.
REFERENCE 1 (bases 1 to 518)
AUTHORS Genoplante.
TITLE Genoplante, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french

plant genomics programme 'Genoplante' (<http://www.genoplante.com>
and <http://genoplante-info.infobiogen.fr>).

FEATURES
 source Location/Qualifiers
 1. .518
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="recital"
 /db_xref="taxon:4565"
 /clone="G550116E20"
 /tissue_type="grain (550 degrees per day after
 pollination)"
 /clone_lib="G550"

ORIGIN

Query Match 29.0%; Score 29.6; DB 14; Length 518;
Best Local Similarity 56.0%; Pred. No. 82;
Matches 56; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
 ||||| | || ||||| | ||| | ||| | | || || ||| |||
Db 393 CTGGTGCCCGCAGTCCTGCACCTCGATGGTGCACGCCTGGTCGAAGCAGATGCAGCACAG 334

Qy 61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGG 100
 |||| || || ||||||| | ||| | ||| |
Db 333 CTCCGTGTGCTCACCTCCGAGAACATGTCGTCGTCGATG 294

RESULT 28

BG606129/c

LOCUS BG606129 564 bp mRNA linear EST 17-APR-2001

DEFINITION WHE2960_H03_006ZS Wheat dormant embryo cDNA library Triticum
aestivum cDNA clone WHE2960_H03_006, mRNA sequence.

ACCESSION BG606129

VERSION BG606129.1 GI:13656112

KEYWORDS EST.

SOURCE Triticum aestivum (bread wheat)

ORGANISM Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.

REFERENCE 1 (bases 1 to 564)

AUTHORS Anderson,O.D., Chao,S., Chin,A., Close,T.J., Doherty,L.,
Fenton,R.D., Lazo,G.R., Rausch,C.J., Walker-Simmons,M.K. and
Wilson,C.

TITLE The structure and function of the expressed portion of the wheat
genomes - Dormant embryo cDNA library

JOURNAL Unpublished (2001)

COMMENT Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oandersn@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.

COMMENT On Jan 10, 2002 this sequence version replaced gi:18120936.
Contact: Waugh R, Marshall DF
Genome Dynamics/Computational Biology
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: est@scri.sari.ac.uk
All sequence has a Phred quality score of 20 or over
Seq primer: M13 reverse.

FEATURES
source Location/Qualifiers
1. .661
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Optic"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="EBem04_SQ003_H10"
/tissue_type="embryo"
/dev_stage="12 DPA"
/lab_host="DH10B"
/clone_lib="embryo, 12 DPA, no treatment, cv Optic, EBem04"
/note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I; Non-normalised library, directionally cloned into pSPORT1. Derived from embryos dissected from developing grains (12 days post anthesis) in glasshouse grown barley plants. Developed as part of the barley transcriptome resources of BBSRC/SEERAD funded cereal IGF (Investigating Gene Function) project."

ORIGIN

Query Match 29.0%; Score 29.6; DB 12; Length 661;
Best Local Similarity 56.0%; Pred. No. 94;
Matches 56; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
||||| | || ||||| | ||| | ||| | ||| ||| |||
Db 453 CTGGTGCCCGCAGTCCCTCCACCTCGATGGTGCACGCCTGGTGAAGCAGATGCAGCACAG 394
Qy 61 ACTGTTGTCACTTTCCGAGGAGACAAGCTGTCCTGGAGG 100
|||| || || ||||| || ||| ||| |
Db 393 CTCCGTGTCGCTCACCTCCGAGAACATGTCGTCGTCGATG 354

RESULT 30
BQ466828/c
LOCUS BQ466828 697 bp mRNA linear EST 30-MAY-2002
DEFINITION HS01L11T HS Hordeum vulgare subsp. vulgare cDNA clone HS01L11
5-PRIME, mRNA sequence.
ACCESSION BQ466828
VERSION BQ466828.1 GI:21274610
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Pooideae; Triticeae; Hordeum.

REFERENCE 1 (bases 1 to 697)

AUTHORS Zhang,H., Potokina,E., Michalek,W., Weschke,W., Stein,N. and Graner,A.

TITLE Barley ESTs from germinating seeds

JOURNAL Unpublished (2002)

COMMENT Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 697 Std Error: 0.00
Plate: 1 row: L column: 11
Seq primer: T3.

FEATURES Location/Qualifiers

source 1. .697
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="HS01L11"
/tissue_type="embryo + scutellum"
/dev_stage="0-16 hours after imbibition"
/lab_host="XL10-Gold"
/clone_lib="HS"
/note="Vector: pBluescript SK+; Site_1: EcoRI (5'-end of cDNA); Site_2: XhoI (3'-end of cDNA); Due to a cloning artefact caused by the kit, in most cases the EcoRI site is NOT present, as well as the EcoRI adapter used for cloning. To excise the insert, restriction sites upstream EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also due to the cloning system used Blue/white selection for recombinants is not 100% reliable."

ORIGIN

Query Match 29.0%; Score 29.6; DB 13; Length 697;
Best Local Similarity 56.0%; Pred. No. 96;
Matches 56; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
||||| | || ||||| | ||| | ||| | ||| ||| |||||
Db 374 CTGGTGCCCGCAGTCCTCCACCTCGATGGTGCACGCCTGGTCGAAGCAGATGCAGCACAG 315

Qy 61 ACTGTTGTCACTTTCGAGGAGAACAAGCTGTCCTGGAGG 100
|||| || || ||||| | ||| | |||
Db 314 CTCCGTGTCGCTCACCTCCGAGAACATGTCGTCGTCGACG 275

RESULT 31
BQ838111/c

LOCUS BQ838111 730 bp mRNA linear EST 08-AUG-2002

DEFINITION WHE2906_F08_K16ZS Wheat aluminum-stressed root tip cDNA library
Triticum aestivum cDNA clone WHE2906_F08_K16, mRNA sequence.

ACCESSION BQ838111

VERSION BQ838111.1 GI:22142429
 KEYWORDS EST.
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooideae; Triticeae; Triticum.
 REFERENCE 1 (bases 1 to 730)
 AUTHORS Anderson,O.D., Chao,S., Chin,A., Close,T.J., Gustafson,J.P.,
 Lazo,G.R., Rausch,C.J., Ross,K., Seaton,C.L. and Wilson,C.
 TITLE The structure and function of the expressed portion of the wheat
 genomes - Aluminum-stressed root tip cDNA library
 JOURNAL Unpublished (2001)
 COMMENT Contact: Olin Anderson
 US Department of Agriculture, Agriculture Research Service, Pacific
 West Area, Western Regional Research Center
 800 Buchanan Street, Albany, CA 94710, USA
 Tel: 5105595773
 Fax: 5105595818
 Email: oandersn@pw.usda.gov
 Sequences have been trimmed to remove vector sequence and low
 quality sequence with phred score less than 20
 Seq primer: SK primer.
 FEATURES Location/Qualifiers
 source 1. .730
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="BH1146"
 /db_xref="taxon:4565"
 /clone="WHE2906_F08_K16"
 /tissue_type="Root tip at 1.0 to 1.5 mm stage"
 /dev_stage="Seedling"
 /lab_host="E. coli SOLR"
 /clone_lib="Wheat aluminum-stressed root tip cDNA library"
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
 Site_1: EcoRI; Site_2: XhoI; Plants were grown under
 hydroponic conditions, root tips were excised and snap
 frozen, total RNA was prepared at University of
 Missouri(Ross, Gustafson). Poly(A) RNA was purified, a
 cDNA library was made, and the cDNA clones were in vivo
 excised to give pBluescript SK- phagemids in the TJ Close
 lab (Chin and Close) at the University of California,
 Riverside. Plasmid DNA preparations and DNA sequencing
 were performed in the OD Anderson lab (all other
 authors)."
 ORIGIN

Query Match 29.0%; Score 29.6; DB 13; Length 730;
 Best Local Similarity 56.0%; Pred. No. 99;
 Matches 56; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
 |||| | || |||| | || | || | || || || || ||
 Db 146 CTGGTGCCCGCAGTCCTGCACCTCGATGGTGCACGCCTGGTCGAAGCAGATGCAGCACAG 87
 Qy 61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGG 100
 |||| || || |||| || | || | || |

Db

86 CTCCGTGTCGCTCACCTCCGAGAACATGTCGTCGTCGATG 47

RESULT 32

CB360743

LOCUS CB360743 429 bp mRNA linear EST 10-NOV-2003

DEFINITION ZF001-P00031-DPE-F-D_B08 GISZF001 Danio rerio cDNA clone
IMAGE:6903233 5' similar to fc20e06.y1 Zebrafish WashU MPIMG EST
Danio rerio cDNA clone IMAGE:3721954 5', mRNA sequence.

ACCESSION CB360743

VERSION CB360743.1 GI:29005688

KEYWORDS EST.

SOURCE Danio rerio (zebrafish)

ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 429)

AUTHORS Mathavan,S., Wei,C., Thoreau,H., Chia,J.M. and Ruan,Y.

TITLE Genome Institute of Singapore, Zebrafish EST Collection

JOURNAL Unpublished (2003)

COMMENT Contact: Ruan Y

Laboratory of Molecular Biotechnology

Genome Institute of Singapore

1 Science Park Road, The Capricorn #05-01, Singapore 117528

Tel: +65 6827 5200

Fax: +65 6827 5201

Email: gisry@nus.edu.sg

GIS Clone ID: ZF001-P00031-PP_D16

PCR Primers

FORWARD: M13

BACKWARD: M13

Plate: ZF001-P00031-DPE-F-D

Seq primer: CCGCATAACTTGATAGCA

High quality sequence stop: 429.

FEATURES

source

Location/Qualifiers

1. .429

/organism="Danio rerio"

/mol_type="mRNA"

/db_xref="taxon:7955"

/clone="IMAGE:6903233"

/tissue_type="Embryo"

/dev_stage="7 Different embryonic Stages(From just
fertilized Embryos to 72 hours just hatched baby fish)"

/lab_host="DH10B"

/clone_lib="GISZF001"

/note="Vector: pDNR-LIB; Site_1: Sfi A (GGCCATTACGGCC);

Site_2: Sfi B (GGCCGCCTCGGCC); Priming method: Sfi-(dT)30

Primed ; Priming sequence: 5.ATTCTAGA GGCCGAGGCGGCC

GACATG(T)30VN ; Directionally cloned, 5' cloning site:

Sfi A site GGCCATTACGGCC ; 5' linker/adaptor sequence:

5.AAGCAGTGGTATCAACGCAGAGTGGCC ; 3' cloning site: Sfi B

site GGCCGCCTCGGCC ; 3' linker/adaptor sequence: same

as the priming sequence ; Average insert size: 2kb ; For

PCR insert analysis: Use M13 Forward and reverse primers ;

Library Amplified Recombinants (inserts): 98% ; Library

complexity: 5x10⁶ ; Full-length construction (method):

SMART, a Clontech method ; Library constructed by: S.
Mathavan, Chia-Lin Wei, and Yijun Ruan Genome Institute of
Singapore"

ORIGIN

Query Match 28.8%; Score 29.4; DB 14; Length 429;
Best Local Similarity 56.8%; Pred. No. 87;
Matches 54; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 6 AGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTGT 65
|| |||| |||||||| || || ||| || || |||||
Db 216 AGAGGAGAGTGAAGACCTCCAGATTGAAGAAACATTCACAGTCAAACATGAAGAGACTGA 275

Qy 66 TGTCACCTTTCCGAGGAGAACAAGCTGTCCTGGAGG 100
| ||||| ||| |||| | | | ||||
Db 276 AGAAGCTTTCAGAGTCAAACATGAAGATCCTGAGG 310

RESULT 33

BQ993297/c

LOCUS BQ993297 463 bp mRNA linear EST 21-AUG-2002

DEFINITION QGF28E04.yg.ab1 QG_EFGHJ lettuce serriola Lactuca sativa cDNA clone
QGF28E04, mRNA sequence.

ACCESSION BQ993297

VERSION BQ993297.1 GI:22412832

KEYWORDS EST.

SOURCE Lactuca sativa

ORGANISM Lactuca sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Cichorioideae;
Cichorieae; Lactuca.

REFERENCE 1 (bases 1 to 463)

AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,
Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
Lai,Z., Church,S., Jackson,L. and Bradford,K.

TITLE Lettuce and Sunflower ESTs from the Compositae Genome Project
<http://compgenomics.ucdavis.edu/>

JOURNAL Unpublished (2002)

COMMENT Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
belongs to contig QG_CA_Contig7941, see <http://cgpdb.ucdavis.edu/>
for details.
Plate: QGF28 row: E column: 04.

FEATURES Location/Qualifiers

source 1..463
/organism="Lactuca sativa"
/mol_type="mRNA"
/cultivar="L.serriola"
/db_xref="taxon:4236"
/clone="QGF28E04"


```

/lab_host="E.coli"
/clone_lib="QG_EFGHJ lettuce serriola"
/note="Vector: pBRcDNASfiAB; The library was constructed
from 10 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_SEQ=Not found"

```

ORIGIN

```

Query Match          28.8%;  Score 29.4;  DB 13;  Length 463;
Best Local Similarity 76.6%;  Pred. No. 90;
Matches   36;  Conservative   0;  Mismatches   11;  Indels       0;  Gaps       0;

```

```

Qy      18  TGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG  64
          ||||| ||||| ||| | ||||| | |||| ||| || ||
Db      256 TGACCACCAGAGTTTGAATTGACCACCGGAGGTGTAGTGGAGCATG 210

```

RESULT 34 CNS02AAN/c

```

LOCUS      CNS02AAN                      1021 bp    DNA        linear    GSS 01-SEP-2000
DEFINITION Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
251G22 of library G from Tetraodon nigroviridis, genomic survey
sequence.

```

```

ACCESSION  AL188312
VERSION     AL188312.1  GI:7826416
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis
ORGANISM    Tetraodon nigroviridis
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
             Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
             Tetradontoidea; Tetraodontidae; Tetraodon.

```

```

REFERENCE   1
AUTHORS     Roest Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
             Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
             Saurin,W. and Weissenbach,J.

```

```

TITLE       Estimate of human gene number provided by genome-wide analysis
            using Tetraodon nigroviridis DNA sequence

```

```

JOURNAL     Nat. Genet. 25 (2), 235-238 (2000)

```

```

MEDLINE     20296633

```

```

PUBMED      10835645

```

```

REFERENCE   2
AUTHORS     Roest Crollius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
             Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
             Saurin,W., Bernot,A. and Weissenbach,J.

```

```

TITLE       Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis

```

```

JOURNAL     Genome Res. 10 (7), 939-949 (2000)

```

```

MEDLINE     20359837

```

```

PUBMED      10899143

```

```

REFERENCE   3  (bases 1 to 1021)

```

AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetraodon>.

FEATURES Location/Qualifiers
source 1. .1021
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="251G22"
/clone_lib="G"
/note="Genoscope sequence ID : C0AG251BD11SP1~end :
PUC-Ori"

ORIGIN

Query Match 28.8%; Score 29.4; DB 29; Length 1021;
Best Local Similarity 60.8%; Pred. No. 1.4e+02;
Matches 48; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
Qy 18 TGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTGTTGTCACCTTCCG 77
|| || |||| ||||| | | |||| |||| |||| ||
Db 173 TGTTCGTGTCGAGTCTTGGACTTCTGCTTCTGGCTGAAGGTCAGATTGTTGCTGCTGATGG 114
Qy 78 AGGAGAACAAGCTGTCCTG 96
||||||| |||||
Db 113 TTCAAACAAGCTCTCCTG 95

RESULT 35

CA628204

LOCUS CA628204 536 bp mRNA linear EST 23-NOV-2002

DEFINITION wle1.pk0005.c7 wle1 Triticum aestivum cDNA clone wle1.pk0005.c7 5'
end, mRNA sequence.

ACCESSION CA628204

VERSION CA628204.1 GI:25206500

KEYWORDS EST.

SOURCE Triticum aestivum (bread wheat)

ORGANISM Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.

REFERENCE 1 (bases 1 to 536)

AUTHORS Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
Miao,G., Caraher,N. and Hanafey,M.K.

TITLE DuPont Wheat cDNA Sequence

JOURNAL Unpublished (2002)

COMMENT Contact: Scott V. Tingey

Crop Genetics

E. I. DuPont de Nemours and Company

1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA

Tel: 302-631-2602

Fax: 302-631-2607

Email: Scott.V.Tingey@USA.dupont.com

Seq primer: M13.

FEATURES
 source Location/Qualifiers
 1. .536
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /db_xref="taxon:4565"
 /clone="wlel.pk0005.c7"
 /tissue_type="leaf"
 /clone_lib="wlel"
 /note="Vector: pBluescript SK+; Site_1: EcoRI; Site_2:
 XhoI; Wheat (Triticum aestivum L.) leaf 7 day old
 etiolated seedling"

ORIGIN

Query Match 28.6%; Score 29.2; DB 14; Length 536;
Best Local Similarity 74.0%; Pred. No. 1.1e+02;
Matches 37; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 2 TGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGT 51
 |||| |||| |||| ||||||||||||| || || || ||
Db 447 TGGTGTCTGAGCTCTCAAACCTCCAGAGTGATGGTCTTGCCGGTGAAGGT 496

RESULT 36

BQ743419/c

LOCUS BQ743419 674 bp mRNA linear EST 17-JUL-2002

DEFINITION WHE4103_G06_N11ZS Wheat salt-stressed root cDNA library Triticum
aestivum cDNA clone WHE4103_G06_N11, mRNA sequence.

ACCESSION BQ743419

VERSION BQ743419.1 GI:21890206

KEYWORDS EST.

SOURCE Triticum aestivum (bread wheat)

ORGANISM Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.

REFERENCE 1 (bases 1 to 674)

AUTHORS Anderson, O.D., Akhunov, E., Chao, S., Crossman, C., Deal, K.,
Dvorak, J., Lazo, G.R., Pham, J., Rausch, C.J., Wilson, C. and Woo, J.

TITLE The structure and function of the expressed portion of the wheat
genomes - Salt-stressed root cDNA library

JOURNAL Unpublished (2002)

COMMENT Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105595773

Fax: 5105595818

Email: oandersn@pw.usda.gov

Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20

Seq primer: SK primer.

FEATURES
 source Location/Qualifiers
 1. .674
 /organism="Triticum aestivum"
 /mol_type="mRNA"

Seq primer: TAATACGACTCACTATAGGG

High quality sequence stop: 436.

FEATURES
source Location/Qualifiers
1. .436
/organism="Prunus persica"
/mol_type="mRNA"
/cultivar="Loring"
/db_xref="taxon:3760"
/clone="PP_LEa0027M13f"
/tissue_type="Mesocarp"
/lab_host="E. coli"
/clone_lib="Peach developing fruit mesocarp"
/note="Vector: pBluescript II SK(-); Site_1: EcoRI;
Site_2: XhoI; authority=Prunus persica L. Batsh; The
sequence has been trimmed to remove vector sequence and
contains a minimum of 100 bases of phred value 20 or
above. For more details on library preparation and
sequence analysis go to
<http://www.genome.clemson.edu/projects/peach>. To order
this clone go to <http://www.genome.clemson.edu/orders>"

ORIGIN

Query Match 28.4%; Score 29; DB 13; Length 436;
Best Local Similarity 58.8%; Pred. No. 1.2e+02;
Matches 50; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 11 AGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTGTTGTCA 70
| | | | | | | | | | | | | | | | | | | | | |
Db 315 ACATTTCCGACGACCAGATTGGCCTTCTTCCCACAGTAGATGAACTGGCCAGTGTAGATA 256

Qy 71 CTTTCCGAGGAGAACAAGCTGTCCT 95
| | | | | | | | | | | | | | | | | | | | | |
Db 255 CCCTCAGCGGCGACGAAGAGCTCGT 231

RESULT 39

BU042469/c

LOCUS BU042469 614 bp mRNA linear EST 26-AUG-2002

DEFINITION PP_LEa0012L15f Peach developing fruit mesocarp Prunus persica cDNA
clone PP_LEa0012L15f, mRNA sequence.

ACCESSION BU042469

VERSION BU042469.1 GI:22482546

KEYWORDS EST.

SOURCE Prunus persica (peach)

ORGANISM Prunus persica

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.

REFERENCE 1 (bases 1 to 614)

AUTHORS Callahan,A., Palmer,M., Main,D., Wing,R. and Abbott,A.

TITLE Peach Model Genome for Rosaceae

JOURNAL Unpublished (2002)

COMMENT Contact: Abbott, A.

Dept of Genetics and Biochemistry

Clemson University

122 Long Hall, Clemson University, Clemson, SC 29634, USA

Tel: 864 656 3060

Fax: 864 656 6879
Email: aalbert@clemson.edu
Total High Quality bases = 498
Seq primer: TAATACGACTCACTATAGGG
High quality sequence stop: 614.

FEATURES
 source
 Location/Qualifiers
 1. .614
 /organism="Prunus persica"
 /mol_type="mRNA"
 /cultivar="Loring"
 /db_xref="taxon:3760"
 /clone="PP_LEa0012L15f"
 /tissue_type="Mesocarp"
 /lab_host="E. coli"
 /clone_lib="Peach developing fruit mesocarp"
 /note="Vector: pBluescript II SK(-); Site_1: EcoRI;
 Site_2: XhoI; authority=Prunus persica L. Batsh; The
 sequence has been trimmed to remove vector sequence and
 contains a minimum of 100 bases of phred value 20 or
 above. For more details on library preparation and
 sequence analysis go to
 <http://www.genome.clemson.edu/projects/peach>. To order
 this clone go to <http://www.genome.clemson.edu/orders>"

ORIGIN

Query Match 28.4%; Score 29; DB 13; Length 614;
Best Local Similarity 58.8%; Pred. No. 1.4e+02;
Matches 50; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 11 AGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTGTTGTCA 70
 | || || || | || || | | | || || || || | | | | | |
Db 330 ACATTTCCGACGACCAGATTGGCTTCTTCCACAGTAGATGAACTGGCCAGTGTAGATA 271

Qy 71 CTTTCCGAGGAGAACAAGCTGTCCT 95
 | || | || || | || | || |
Db 270 CCCTCAGCGGCGACGAAGAGCTCGT 246

RESULT 40

BU046581/c

LOCUS BU046581 630 bp mRNA linear EST 26-AUG-2002

DEFINITION PP_LEa0026M12f Peach developing fruit mesocarp Prunus persica cDNA
clone PP_LEa0026M12f, mRNA sequence.

ACCESSION BU046581

VERSION BU046581.1 GI:22486658

KEYWORDS EST.

SOURCE Prunus persica (peach)

ORGANISM Prunus persica

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.

REFERENCE 1 (bases 1 to 630)

AUTHORS Callahan,A., Palmer,M., Main,D., Wing,R. and Abbott,A.

TITLE Peach Model Genome for Rosaceae

JOURNAL Unpublished (2002)

COMMENT Contact: Abbott, A.

Dept of Genetics and Biochemistry

Clemson University
122 Long Hall, Clemson University, Clemson, SC 29634, USA
Tel: 864 656 3060
Fax: 864 656 6879
Email: aalbert@clemson.edu
Total High Quality bases = 523
Seq primer: TAATACGACTCACTATAGGG
High quality sequence stop: 630.

FEATURES
 source Location/Qualifiers
 1. .630
 /organism="Prunus persica"
 /mol_type="mRNA"
 /cultivar="Loring"
 /db_xref="taxon:3760"
 /clone="PP_LEa0026M12f"
 /tissue_type="Mesocarp"
 /lab_host="E. coli"
 /clone_lib="Peach developing fruit mesocarp"
 /note="Vector: pBluescript II SK(-); Site_1: EcoRI;
 Site_2: XhoI; authority=Prunus persica L. Batsh; The
 sequence has been trimmed to remove vector sequence and
 contains a minimum of 100 bases of phred value 20 or
 above. For more details on library preparation and
 sequence analysis go to
 <http://www.genome.clemson.edu/projects/peach>. To order
 this clone go to <http://www.genome.clemson.edu/orders>"

ORIGIN

Query Match 28.4%; Score 29; DB 13; Length 630;
Best Local Similarity 58.8%; Pred. No. 1.4e+02;
Matches 50; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 11 AGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTGTTGTCA 70
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 332 ACATTTCCGACGACCAAGATTGGCCTTCTCCACAGTAGATGAACTGGCCAGTGTAGATA 273

Qy 71 CTTTCCGAGGAGAACAAGCTGTCCT 95
 | | | | | | | | | | | | | | | | | | | | | |
Db 272 CCCTCAGCGGCGACGAAGAGCTCGT 248

RESULT 41

BU044321/c

LOCUS BU044321 635 bp mRNA linear EST 26-AUG-2002

DEFINITION PP_LEa0018018f Peach developing fruit mesocarp Prunus persica cDNA
clone PP_LEa0018018f, mRNA sequence.

ACCESSION BU044321

VERSION BU044321.1 GI:22484398

KEYWORDS EST.

SOURCE Prunus persica (peach)

ORGANISM Prunus persica

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.

REFERENCE 1 (bases 1 to 635)

AUTHORS Callahan,A., Palmer,M., Main,D., Wing,R. and Abbott,A.

TITLE Peach Model Genome for Rosaceae

JOURNAL Unpublished (2002)
COMMENT Contact: Abbott, A.
Dept of Genetics and Biochemistry
Clemson University
122 Long Hall, Clemson University, Clemson, SC 29634, USA
Tel: 864 656 3060
Fax: 864 656 6879
Email: aalbert@clemson.edu
Total High Quality bases = 522
Seq primer: TAATACGACTCACTATAGGG
High quality sequence stop: 635.

FEATURES
source Location/Qualifiers
1. .635
/organism="Prunus persica"
/mol_type="mRNA"
/cultivar="Loring"
/db_xref="taxon:3760"
/clone="PP_LEa0018018f"
/tissue_type="Mesocarp"
/lab_host="E. coli"
/clone_lib="Peach developing fruit mesocarp"
/note="Vector: pBluescript II SK(-); Site_1: EcoRI;
Site_2: XhoI; authority=Prunus persica L. Batsh; The
sequence has been trimmed to remove vector sequence and
contains a minimum of 100 bases of phred value 20 or
above. For more details on library preparation and
sequence analysis go to
<http://www.genome.clemson.edu/projects/peach>. To order
this clone go to <http://www.genome.clemson.edu/orders>"

ORIGIN

Query Match 28.4%; Score 29; DB 13; Length 635;
Best Local Similarity 58.8%; Pred. No. 1.4e+02;
Matches 50; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 11 AGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTGTTGTCA 70
| | | | | | | | | | | | | | | | | | | | | |
Db 315 ACATTTCCGACGACCAGATTGGCCTTCTTCCCACAGTAGATGAACTGGCCAGTGTAGATA 256

Qy 71 CTTTCCGAGGAGAACAAGCTGTCCT 95
| | | | | | | | | | | | | | | | | | | | | |
Db 255 CCCTCAGCGGCGACGAAGAGCTCGT 231

RESULT 42
BH109216/c
LOCUS BH109216 735 bp DNA linear GSS 19-JUL-2001
DEFINITION RPCI-24-340C23.TJ RPCI-24 Mus musculus genomic clone
RPCI-24-340C23, genomic survey sequence.
ACCESSION BH109216
VERSION BH109216.1 GI:14942075
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 735)

AUTHORS Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
Russell,D., de Jong,P. and Fraser,C.M.

TITLE Mouse BAC End Sequences from Library RPCI-24

JOURNAL Unpublished (1999)

COMMENT Other_GSSs: RPCI-24-340C23.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC
library availability, please contact Pieter de Jong
(pdejong@mail.cho.org). Clones may be purchased from BACPAC
Resources (<http://www.chori.org/bacpac/orderingframe.htm>). BAC end
page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 340 row: C column: 23
Seq primer: SP6
Class: BAC ends.

FEATURES Location/Qualifiers

source 1. .735
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-340C23"
/sex="Male"
/cell_type="Spleen/Brain"
/clone_lib="RPCI-24"
/note="Vector: pTARBAC1; Site_1: BamH1; Site_2: BamH1;
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pTARBAC1 cloning vector at the
BamH1 sites using MboI partially digested male C57BL/6J
DNA."

ORIGIN

Query Match 28.4%; Score 29; DB 28; Length 735;
Best Local Similarity 63.8%; Pred. No. 1.5e+02;
Matches 44; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 27 GAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTGTTGTCACTTTCCGAGGAGAACA 86
|| | ||| ||||| || ||| | | |||| | | |||||

Db 666 GATTTCTGGAGCTTCCACTGTCTGTCAAGTTGTGGCACATGTCAGCTCACAAGGAGAACA 607

Qy 87 AGCTGTCCT 95
| ||| |

Db 606 AACTGGCTT 598

RESULT 43 AI117880/c

LOCUS AI117880 342 bp mRNA linear EST 02-SEP-1998
DEFINITION uc41f02.r1 Soares_mammary_gland_NMLMG Mus musculus cdna clone
IMAGE:1400571 5', mRNA sequence.
ACCESSION AI117880

```

VERSION      AI117880.1  GI:3518204
KEYWORDS     EST.
SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1 (bases 1 to 342)
AUTHORS      Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
              Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
              Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
              Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
              Waterston,R.
TITLE        The WashU-HHMI Mouse EST Project
JOURNAL      Unpublished (1996)
COMMENT      Contact: Marra M/Mouse EST Project
              WashU-HHMI Mouse EST Project
              Washington University School of MedicineP
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: mouseest@watson.wustl.edu
              This clone is available royalty-free through LLNL ; contact the
              IMAGE Consortium (info@image.llnl.gov) for further information.
              MGI:912287
              Seq primer: -28ml3 rev2 ET from Amersham
              High quality sequence stop: 297.
FEATURES
  source      Location/Qualifiers
              1. .342
                /organism="Mus musculus"
                /mol_type="mRNA"
                /db_xref="taxon:10090"
                /clone="IMAGE:1400571"
                /sex="female (lactating)"
                /tissue_type="mammary gland"
                /lab_host="DH10B"
                /clone_lib="Soares_mammary_gland_NMLMG"
                /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                polylinker; 1st strand cDNA was prepared from mammary
                gland tissue from a lactating female, and was then primed
                with a Not I - oligo(dT) primer. Double-stranded cDNA was
                ligated to Eco RI adaptors (Pharmacia), digested with Not
                I and cloned into the Not I and Eco RI sites of the
                modified pT7T3 vector. Library is normalized. Library
                was constructed by Bento Soares and M. Fatima Bonaldo. "
ORIGIN
Query Match      28.2%; Score 28.8; DB 9; Length 342;
Best Local Similarity 58.0%; Pred. No. 1.2e+02;
Matches 51; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy      14 TCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTGTTGTCACCTT 73
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      260 TTTCTCAACTATAGAATCTAGTTGTGAAGACTTTTCATTAAGTTGCTCTTGAGAACACTT 201

Qy      74 TCCGAGGAGAACAAGCTGTCCTGGAGGC 101
      | ||| |||| ||| |||| ||| ||| |||
Db      200 TTCGATGAGAGCGATCTGTTCTTGTAGC 173

```

RESULT 44
 AA177634/c
 LOCUS AA177634 398 bp mRNA linear EST 16-FEB-1997
 DEFINITION mt32h12.r1 Soares mouse 3NbMS Mus musculus cDNA clone IMAGE:622823
 5', mRNA sequence.
 ACCESSION AA177634
 VERSION AA177634.1 GI:1758868
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 398)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 TITLE The WashU-HHMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HHMI Mouse EST Project
 Washington University School of MedicineP
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:383647
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 371.
 FEATURES Location/Qualifiers
 source 1. .398
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:622823"
 /sex="male"
 /tissue_type="Spleen"
 /dev_stage="4 weeks"
 /lab_host="DH10B"
 /clone_lib="Soares mouse 3NbMS"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5'
 TGTACCAATCTGAAGTGGGAGCGGCCGCGCTGTTTTTTTTTTTTTTTTTTTTTTT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT7T3 vector. RNA
 provided by Dr. Bertrand Jordan. Library went through
 three rounds of normalization, and was constructed by
 Bento Soares and M.Fatima Bonaldo."
 ORIGIN

Query Match 28.2%; Score 28.8; DB 9; Length 398;
Best Local Similarity 58.0%; Pred. No. 1.3e+02;
Matches 51; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 14 TCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTGTTGTCACTT 73
| ||| | || ||| | | | ||| ||| | | ||| |||
Db 240 TTCTCAACTATAGAATCTAGTTGTGAAGACTTTTCATTAAGTTGCTCTTGAGAACACTT 181

Qy 74 TCCGAGGAGAACAAAGCTGTCCTGGAGGC 101
| ||| |||| | | |||| || | ||
Db 180 TTCGATGAGAGCGATCTGTTCTTGTAGC 153

RESULT 45
BG550348/c

LOCUS BG550348 416 bp mRNA linear EST 05-APR-2001
DEFINITION 947039G04.x2 947 - 2 week shoot from Barkan lab Zea mays cDNA, mRNA
sequence.

ACCESSION BG550348
VERSION BG550348.1 GI:13558993
KEYWORDS EST.

SOURCE Zea mays
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 416)

AUTHORS Walbot,V.

TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
University

JOURNAL Unpublished (1999)

COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 947039 row: G column: 04.

FEATURES Location/Qualifiers

source

1. .416
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="B73"
/db_xref="taxon:4577"
/tissue_type="leaf and stem, including leaf base"
/dev_stage="2 week old seedling (3 leaves)"
/lab_host="XL1-Blue"
/clone_lib="947 - 2 week shoot from Barkan lab"
/note="Organ: shoot; Vector: Lambda ZAP (pBlueScript SK-);
Site_1: EcoRI; Site_2: XhoI; Directionally cloned using
Stratagene's UniZap XR cDNA cloning kit with the 5' end
at the EcoRI site. The library represents 8 x 10e5
independent recombinant phage. The plants were greenhouse
grown."

ORIGIN

Query Match 28.2%; Score 28.8; DB 12; Length 416;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 48; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
|| | ||| | | || | || || || | | || | || |

Db 105 CTATCAAGTGTGTAGTGTGTCTTCGAGAAGTTTGTAGAGCCTACTGCTGCTGCTGTATAT 46

Qy 61 ACTGTTGTCACTTTCCGAGG 80

|||| | || ||| || |||

Db 45 ACTGATATCGCTTGCCAAGG 26

RESULT 46

BQ557757

LOCUS BQ557757 510 bp mRNA linear EST 20-JUN-2002

DEFINITION H4048B01-3 NIA Mouse 7.4K cDNA Clone Set Mus musculus cDNA clone
H4048B01 3', mRNA sequence.

ACCESSION BQ557757

VERSION BQ557757.1 GI:21458642

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 510)

AUTHORS VanBuren,V., Piao,Y., Dudekula,D.B., Qian,Y., Carter,M.G.,
Martin,P.R., Stagg,C.A., Bassey,U., Aiba,K., Hamatani,T.,
Kargul,G.J., Luo,A.G., Kelso,J., Hide,W. and Ko,M.S.H.

TITLE Assembly, verification, and initial annotation of NIA 7.4K mouse
cDNA clone set

JOURNAL Genome Res. 12 (12), 1999-2003 (2002)

MEDLINE 22354164

PUBMED 12466305

COMMENT Other_ESTs: H4048B01-5

Contact: Yong Qian

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA

Email: cdna@lgsun.grc.nia.nih.gov

This clone set has been freely distributed to the community. Please
visit http://lgsun.grc.nia.nih.gov/cDNA/NIA_7_4k.html for details.

Plate: H4048 row: B column: 01

Seq primer: -21M13 Forward

High quality sequence stop: 510

POLYA=Yes.

FEATURES Location/Qualifiers

source

1..510

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="niaEST:H4048B01-3"

/db_xref="taxon:10090"

/clone="H4048B01"

/sex="mixed"

/dev_stage="mixed"

/lab_host="DH10B"
/clone_lib="NIA Mouse 7.4K cDNA Clone Set"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
clone is among a rearranged set of 7,407 clones from more
than 20 cDNA libraries."

ORIGIN

Query Match 28.2%; Score 28.8; DB 13; Length 510;
Best Local Similarity 58.0%; Pred. No. 1.5e+02;
Matches 51; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 14 TCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTGTTGTCACCTT 73
| ||| | || ||| | | | ||| ||| | | ||| |||
Db 150 TTTCTCAACTATAGAATCTAGTTGTGAAGACTTTTCATTAAGTTGCTCTTGAGAACACTT 209

Qy 74 TCCGAGGAGAACAAGCTGTCCTGGAGGC 101
| ||| |||| | | |||| || | ||
Db 210 TTCGATGAGAGCGATCTGTTCTTGTAGC 237

RESULT 47

BX514645/c

LOCUS BX514645 524 bp mRNA linear EST 25-JUN-2003

DEFINITION BX514645 Soares mouse 3NbMS Mus musculus cDNA clone IMAGp952C2329 ;
IMAGE:622823, mRNA sequence.

ACCESSION BX514645

VERSION BX514645.1 GI:32244604

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 524)

AUTHORS Heil,O., Ebert,L., Neubert,P., Peters,M., Radelof,U., Schneider,D.
and Korn,B.

TITLE Mouse UnigeneSet - RZPD2

JOURNAL Unpublished (2003)

COMMENT Contact: Ina Rolfs

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany

RZPD; IMAGp952C2329.

RZPDLIB; I.M.A.G.E. cDNA Clone Collection;

Mouse UnigeneSet - RZPD2 (RZPDLIB No.981)

[http://www.rzpd.de/CloneCards/cgi-](http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=981)

[bin/showLib.pl.cgi/response?libNo=981](http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=981) Contact: Ina Rolfs

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

Heubnerweg 6, D-14059 Berlin, Germany

Tel: +49 30 32639 101

Fax: +49 30 32639 111

www.rzpd.de

This clone is available royalty-free from RZPD;

contact RZPD (clone@rzpd.de) for further information. Seq primer:

T7, Primer sequence: TAATACGACTCACTATAGGG.

FEATURES

source

Location/Qualifiers

1..524

/organism="Mus musculus"

/mol_type="mRNA"

```

/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGp952C2329 ; IMAGE:622823"
/sex="male"
/tissue_type="Spleen"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares mouse 3NbMS"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTGTTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. RNA
provided by Dr. Bertrand Jordan. Library went through
three rounds of normalization, and was constructed by
Bento Soares and M.Fatima Bonaldo."

```

ORIGIN

```

Query Match          28.2%; Score 28.8; DB 13; Length 524;
Best Local Similarity 58.0%; Pred. No. 1.5e+02;
Matches 51; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```

```

Qy      14 TCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTGTTGTCACTT 73
        | | | | | | | | | | | | | | | | | | | | | | | |
Db      238 TTTCTCAACTATAGAATCTAGTTGTGAAGACTTTTCATTAAGTTGCTCTTGAGAACACTT 179

Qy      74 TCCGAGGAGAACAAGCTGTCCTGGAGGC 101
        | | | | | | | | | | | | | | | | | |
Db      178 TTCGATGAGAGCGATCTGTTCTTGTAGC 151

```

RESULT 48

BX520764/c

LOCUS BX520764 536 bp mRNA linear EST 27-JUN-2003

DEFINITION BX520764 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
IMAGp998K043537 ; IMAGE:1400571, mRNA sequence.

ACCESSION BX520764

VERSION BX520764.1 GI:32301442

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 536)

AUTHORS Heil,O., Ebert,L., Neubert,P., Peters,M., Radelof,U., Schneider,D.
and Korn,B.

TITLE Mouse UnigeneSet - RZPD2

JOURNAL Unpublished (2003)

COMMENT Contact: Ina Rolfs

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany

RZPD; IMAGp998K043537.

RZPDLIB; I.M.A.G.E. cDNA Clone Collection;

Mouse UnigeneSet - RZPD2 (RZPDLIB No.981)

<http://www.rzpd.de/CloneCards/cgi->

bin/showLib.pl.cgi/response?libNo=981 Contact: Ina Rolfs
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Heubnerweg 6, D-14059 Berlin, Germany
 Tel: +49 30 32639 101
 Fax: +49 30 32639 111
 www.rzpd.de
 This clone is available royalty-free from RZPD;
 contact RZPD (clone@rzpd.de) for further information. Seq primer:
 T7, Primer sequence: TAATACGACTCACTATAGGG.

FEATURES
 source Location/Qualifiers
 1. .536
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGp998K043537 ; IMAGE:1400571"
 /sex="female (lactating)"
 /tissue_type="mammary gland"
 /lab_host="DH10B"
 /clone_lib="Soares_mammary_gland_NMLMG"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; 1st strand cDNA was prepared from mammary
 gland tissue from a lactating female, and was then primed
 with a Not I - oligo(dT) primer. Double-stranded cDNA was
 ligated to Eco RI adaptors (Pharmacia), digested with Not
 I and cloned into the Not I and Eco RI sites of the
 modified pT7T3 vector. Library is normalized. Library
 was constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 28.2%; Score 28.8; DB 13; Length 536;
 Best Local Similarity 58.0%; Pred. No. 1.5e+02;
 Matches 51; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 14 TCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTGTTGTCACCTT 73
 | | | | | | | | | | | | | | | | | | | | | |
 Db 271 TTTCTCAACTATAGAATCTAGTTGTGAAGACTTTTCATTAAGTTGCTCTTGAGAACACTT 212
 Qy 74 TCCGAGGAGAACAAGCTGTCCTGGAGGC 101
 | | | | | | | | | | | | | | | | | |
 Db 211 TTCGATGAGAGCGATCTGTTCTTGTAGC 184

RESULT 49

AI591944/c

LOCUS AI591944 598 bp mRNA linear EST 15-MAR-2000
 DEFINITION mt32h12.y1 Soares mouse 3NbMS Mus musculus cDNA clone IMAGE:622823
 5', mRNA sequence.

ACCESSION AI591944

VERSION AI591944.1 GI:4600992

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 598)

AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
 Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,

Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.

TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
This read is a RESEQUENCE of a previously sequenced mouse clone
This read has been verified (found to hit its original self in the
correct orientation)
Putative full length read
vector to vector length is 915
MGI:383647
Seq primer: -40RP from Gibco
High quality sequence stop: 460
POLYA=No.

FEATURES Location/Qualifiers
source 1. .598
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:622823"
/sex="male"
/tissue_type="Spleen"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares mouse 3NbMS"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTGTTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. RNA
provided by Dr. Bertrand Jordan. Library went through
three rounds of normalization, and was constructed by
Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 28.2%; Score 28.8; DB 9; Length 598;
Best Local Similarity 58.0%; Pred. No. 1.6e+02;
Matches 51; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 14 TCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTGTTGTCACTT 73
| ||| | || ||| | | ||| ||| | | |||| | | ||||
Db 240 TTTCTCAACTATAGAATCTAGTTGTGAAGACTTTTCATTAAGTTGCTCTTGAGAACACTT 181
Qy 74 TCCGAGGAGACAAGCTGTCCTGGAGGC 101
| ||| |||| | | |||| || | ||
Db 180 TTCGATGAGAGCGATCTGTTCTTGTAGC 153

RESULT 50
 DR36H15S/c
 LOCUS DR36H15S 654 bp DNA linear GSS 22-NOV-2002
 DEFINITION Danio rerio genomic clone DKEY-36H15, genomic survey sequence.
 ACCESSION AL987137
 VERSION AL987137.1 GI:25180574
 KEYWORDS GSS.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.
 REFERENCE 1 (bases 1 to 654)
 AUTHORS Humphray,S.J., Huckle,E. and Hunt,S.E.
 TITLE Direct Submission
 JOURNAL Submitted (14-NOV-2002) The Sanger Institute, Wellcome Trust Genome
 Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
 humquery@sanger.ac.uk Unpublished
 COMMENT This sequence was generated from the SP6 end of BAC 36H15. 36H15 is
 part of the Daniokey BAC Library created by R. Plasterk and N.V.
 Keygene.
 Further details: http://www.sanger.ac.uk/Projects/D_rerio/.
 FEATURES Location/Qualifiers
 source 1..654
 /organism="Danio rerio"
 /mol_type="genomic DNA"
 /db_xref="taxon:7955"
 /clone="DKEY-36H15"
 /tissue_type="Testis"
 /note="vector pIndigoBAC-536"

ORIGIN

Query Match 28.2%; Score 28.8; DB 29; Length 654;
 Best Local Similarity 59.3%; Pred. No. 1.7e+02;
 Matches 48; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 9 TGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTGTTGT 68
 || |||| | | | |||| | | |||| | | |||||
 Db 90 TGTGATCGTAGTAGTGCTGTGTCTGTTGAAATTTCAAGGTGACGTNCTGACTGTTGG 31

Qy 69 CACTTTCCGAGGAGAACAAGC 89
 | | | |||| | ||| |
 Db 30 AAGATGCTGAGGCCAGCAAAC 10

Search completed: April 29, 2004, 18:39:20
 Job time : 338.622 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 29, 2004, 14:53:09 ; Search time 435.147 Seconds
(without alignments)
10159.758 Million cell updates/sec

Title: US-09-989-981A-9_COPY_3_104
Perfect score: 102
Sequence: 1 ctggtaggtgagatctctga.....aacaagctgtcctggaggcc 102

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*

28: em_un:*
 29: em_vi:*
 30: em_htg_hum:*
 31: em_htg_inv:*
 32: em_htg_other:*
 33: em_htg_mus:*
 34: em_htg_pln:*
 35: em_htg_rod:*
 36: em_htg_mam:*
 37: em_htg_vrt:*
 38: em_sy:*
 39: em_htgo_hum:*
 40: em_htgo_mus:*
 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		Query					Description
No.	Score	Match	Length	DB	ID		
c	1	102	100.0	1017	10	F351799S02	AF351800 Mus muscu
c	2	102	100.0	2019	6	AX685731	AX685731 Sequence
c	3	102	100.0	2284	10	AY196216	AY196216 Mus muscu
c	4	102	100.0	3674	10	AF324495	AF324495 Mus muscu
	5	102	100.0	6043	6	AX685737	AX685737 Sequence
c	6	100.4	98.4	2285	10	AY196215	AY196215 Mus muscu
c	7	94.6	92.7	4829	10	AF351785	AF351785 Rattus no
	8	94.6	92.7	40929	10	AY145899	AY145899 Rattus no
c	9	94.6	92.7	237445	2	AC120701	AC120701 Rattus no
	10	94.6	92.7	312858	2	AC112747	AC112747 Rattus no
c	11	87.6	85.9	2022	9	AF320294	AF320294 Homo sapi
c	12	87.6	85.9	2669	6	AX685735	AX685735 Sequence
c	13	87.6	85.9	139342	9	AC108476	AC108476 Homo sapi
	14	87.6	85.9	159346	2	AC145533	AC145533 Lemur cat
c	15	86	84.3	2679	9	AF324494	AF324494 Homo sapi
c	16	86	84.3	4665	9	F351812S02	AF351813 Homo sapi
c	17	86	84.3	127066	9	AC084265	AC084265 Homo sapi
c	18	84.4	82.7	178016	2	AC146787	AC146787 Aotus nan
c	19	84.4	82.7	185045	2	AC146466	AC146466 Callithri
	20	82.8	81.2	202533	2	AC146464	AC146464 Saimiri s
c	21	82.8	81.2	207760	2	AC146286	AC146286 Callicebu
c	22	52.4	51.4	135280	2	AC146282	AC146282 Takifugu
	23	35.8	35.1	169570	5	AL928999	AL928999 Zebrafish
c	24	34.2	33.5	190952	5	BX004832	BX004832 Zebrafish
	25	34.2	33.5	226929	2	BX571838	BX571838 Danio rer
	26	32.2	31.6	203	9	HUMCFTR10	M55034 Human cysti
	27	32.2	31.6	206	9	HUMCFTR1	M55025 Homo sapien
	28	32.2	31.6	261	9	MFCFTRW11	AF162161 Macaca fa
	29	32.2	31.6	261	9	MFUSCFTR11	AF162357 Macaca fu
	30	32.2	31.6	261	9	MNCFTR11	AF162384 Macaca ne
	31	32.2	31.6	261	9	PHACFTR11	AF162411 Papio ham
	32	32.2	31.6	261	9	RMCFTR11	AF016934 Macaca mu
	33	32.2	31.6	420	6	AR166291	AR166291 Sequence

34	32.2	31.6	795	6	AR381208	AR381208 Sequence
35	32.2	31.6	831	9	HUMCFTRA10	M55115 Human cysti
36	32.2	31.6	831	11	G18240	G18240 sWSS853 Eri
37	32.2	31.6	2640	6	AR076451	AR076451 Sequence
38	32.2	31.6	2640	6	I46970	I46970 Sequence 1
39	32.2	31.6	2908	9	HUMCFTR10E	L49160 Homo sapien
40	32.2	31.6	4443	6	AR240920	AR240920 Sequence
41	32.2	31.6	4443	6	AR240921	AR240921 Sequence
42	32.2	31.6	4443	6	AR240922	AR240922 Sequence
43	32.2	31.6	4443	6	AR240923	AR240923 Sequence
44	32.2	31.6	4443	6	AR240924	AR240924 Sequence
45	32.2	31.6	4443	6	AR240925	AR240925 Sequence
46	32.2	31.6	4443	6	AR240926	AR240926 Sequence
47	32.2	31.6	4443	6	AR240927	AR240927 Sequence
48	32.2	31.6	4443	6	AR240928	AR240928 Sequence
49	32.2	31.6	4443	6	AR240929	AR240929 Sequence
50	32.2	31.6	4443	6	AX111569	AX111569 Sequence

ALIGNMENTS

RESULT 1

F351799S02/c

LOCUS F351799S02 1017 bp DNA linear ROD 23-AUG-2002

DEFINITION Mus musculus sterolin 2 (Abcg8) gene, exon 2.

ACCESSION AF351800

VERSION AF351800.1 GI:18996438

KEYWORDS .

SEGMENT 2 of 13

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1017)

AUTHORS Lu,K., Lee,M.-H., Yu,H., Zhou,Y., Sandell,S.A., Salen,G. and
Patel,S.B.

TITLE Molecular cloning, genomic organization, genetic variations, and
characterization of murine sterolin genes Abcg5 and Abcg8

JOURNAL J. Lipid Res. 43 (4), 565-578 (2002)

MEDLINE 21904563

PUBMED 11907139

REFERENCE 2 (bases 1 to 1017)

AUTHORS Lu,K., Zhou,Y., Lee,M.-H. and Patel,S.B.

TITLE Direct Submission

JOURNAL Submitted (21-FEB-2001) Division of Endocrinology, Diabetes and
Medical Genetics, Medical University of South Carolina, 114 Doughty
St., STB 541, Charleston, SC 29403, USA

FEATURES Location/Qualifiers

source

1..1017

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="129/Sv"

/db_xref="taxon:10090"

/chromosome="17"

/map="between Mit41 and Mit189"

/clone="329B11"

exon 206. .310
/gene="Abcg8"
/number=2

ORIGIN

Query Match 100.0%; Score 102; DB 10; Length 1017;
Best Local Similarity 100.0%; Pred. No. 1.1e-23;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
|||||
Db 310 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 251

Qy 61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 102
|||||
Db 250 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 209

RESULT 2

AX685731/c

LOCUS AX685731 2019 bp DNA linear PAT 29-MAR-2003

DEFINITION Sequence 3 from Patent WO02081691.

ACCESSION AX685731

VERSION AX685731.1 GI:29371740

KEYWORDS .

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

AUTHORS Hobbs,H.H., Shan,B., Barnes,R. and Tian,H.

TITLE Abcg5 and abcg8: compositions and methods of use

JOURNAL Patent: WO 02081691-A 3 17-OCT-2002;
Tularik Inc. (US) ; BOARD OF REGENTS UNIVERSITY OF TEXAS SYSTEM
(US)

FEATURES Location/Qualifiers

source 1. .2019
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

CDS 1. .2019
/note="unnamed protein product; mouse ABCG8 (mABCG8)"
/codon_start=1
/protein_id="CAD86571.1"
/db_xref="GI:29371741"
/db_xref="REMTREMBL:CAD86571"
/translation="MAEKTKEETQLWNGTVLQDASGLQDSLFSSSEDNSLYFTYSGQS
NTLEVRDLTYQVDIASQVPWFQLAQFKIPWRSHSSQDSCELGIRNLSFKVRSQGMLA
IIGSSGCGRASLLDVITGRGHGGMKSGQIWINQGPSTPQLVRKCVAVHRQHDQLLPN
LTVRETLAFIAQMRLPRTFSQAQRDKRVEDVIAELRLRQCANTRVGNTYVRGVSGGER
RRVSIGVQLLWNPGLILDEPTSGLDSFTAHLNLTLSRLAKGNRLVLISLHQPRSDI
FRLFDLVLLMTSGTPIIYLGAAQQMVQYFTSIGHPCPRYSNPADFYVDLTSIDRRSKER
EVATVEKAQSLAALFLEKVQGFDDFLWKAEAKELNTSTHTVSLTLTQDTCGTAVELP
GMIEQFSTLIRROI SNDFRDLPTLLIHGSEACIMSLIIGFLYYGHGAKQLSFMDTAAL
LFMIGALIPFNVIDVSKCHSERSMLYYELEDGLYTAGPYFFAKILGELPEHCAYVI
IYAMPIYWLTLNLRPVPELFLHFLLVVVFCCRTMALAASAMLP TFHMSSFFCNALY
NSFYLTAGFMINLDNLWIVPAWISKLSFLRWCFSGLMQIQFNGHLYTTQIGNFTFSIL

GDTMISAMDLNSHPLYAIYLIVIGISYGFLFLYYLSLKLKLIKQKSIQDW"

ORIGIN

Query Match 100.0%; Score 102; DB 6; Length 2019;
Best Local Similarity 100.0%; Pred. No. 1.1e-23;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
|||||
Db 165 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 106

Qy 61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 102
|||||
Db 105 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 64

RESULT 3

AY196216/c

LOCUS AY196216 2284 bp mRNA linear ROD 01-JUN-2003

DEFINITION Mus musculus strain PERA/Ei ATP-binding cassette sub-family G
member 8 (Abcg8) mRNA, complete cds.

ACCESSION AY196216

VERSION AY196216.1 GI:31322261

KEYWORDS .

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2284)

AUTHORS Wittenburg,H., Lyons,M.A., Li,R., Churchill,G.A., Carey,M.C. and
Paigen,B.

TITLE Primary Roles of FXR and ABCG5/ABCG8 in Cholesterol Gallstone
Susceptibility: Evidence from a Cross of PERA/Ei and I/Ln Inbred
Mice

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2284)

AUTHORS Lyons,M.A., Wittenburg,H., Walsh,K.A., Carey,M.C. and Paigen,B.

TITLE Direct Submission

JOURNAL Submitted (12-DEC-2002) The Jackson Laboratory, 600 Main Street,
Bar Harbor, ME 04609, USA

FEATURES Location/Qualifiers

source 1. .2284

/organism="Mus musculus"

/mol_type="mRNA"

/strain="PERA/Ei"

/db_xref="taxon:10090"

/chromosome="17"

/map="55 cM"

/sex="male"

/tissue_type="liver"

gene 1. .2284

/gene="Abcg8"

CDS 102. .2120

/gene="Abcg8"

/note="ATP-dependent canalicular cholesterol transporter;
white subfamily"

/codon_start=1

/product="ATP-binding cassette sub-family G member 8"
 /protein_id="AA045096.1"
 /db_xref="GI:31322262"
 /translation="MAEKTKEETQLWNGTVLQDASGLQDSLFSSES DNSLYFTTSGQS
 NTLEVRDLTYQVDIASQVPWFELAQFKIPWRSHSSQDSCELGIRNLSEFKVRSGQMLA
 IIGSSGCGRASLLDVITGRGHGGMKSGQIWINQGPSTPQLVRKCVAVHRQHDQLLPN
 LTVRETLAFIAQMRLPRTFSQAQRDKRVEDVIAELRLRQCANTRVGNTYVRGVSGGER
 RRVSIGVQLLWNP GILILDEPTSGLD SFTAHLNLTLSRLAKGNRLVLISLHQPRSDI
 FRLFDLVLLMTSGTPIYLGAAQOMVQYFTSIGHPCPRYSNPADFYVDLTSIDRRSKER
 EVATVEKAQSLAALFLEKVGQGFDDFLWKAEAKELNTSTHTVSLTLTQDTCGTAVELP
 GMIEQFSTLIRRIISNDFRDLPTLLIHGSEACLSLIIGFLYYGHGAKQLSFMDTAAL
 LFMIGALIPFNVILDVVS KCHSERSMLYIELEDGLYTAGPYFFAKILGELPEHCAYVI
 IYAMPIYWLTNLRPVPELFLHFLLVVLFVFCRTMALAASAMLPTEHMS SFFCNALY
 NSFYLTAGFMINLDNLWIVPAWISKLSFLRWCFSGLMQIQFNHLYTTQIGNFTFSIL
 GDTMISAMDNLN SHPLYAIYLIVIGISYGFLFLYYLSLKLKQKSIQDW"

ORIGIN

Query Match 100.0%; Score 102; DB 10; Length 2284;
 Best Local Similarity 100.0%; Pred. No. 1.1e-23;
 Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 266 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 207

 Qy 61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 102
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 206 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 165

RESULT 4

AF324495/c

LOCUS AF324495 3674 bp mRNA linear ROD 07-AUG-2001

DEFINITION Mus musculus sterolin-2 (Abcg8) mRNA, complete cds.

ACCESSION AF324495

VERSION AF324495.1 GI:15088541

KEYWORDS .

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 3674)

AUTHORS Lu,K., Lee,M.H., Hazard,S., Brooks-Wilson,A., Hidaka,H., Kojima,H.,
 Ose,L., Stalenhoef,A.F., Mietinnen,T., Bjorkhem,I., Bruckert,E.,
 Pandya,A., Brewer,H.B. Jr., Salen,G., Dean,M., Srivastava,A. and
 Patel,S.B.

TITLE Two genes that map to the STSL locus cause sitosterolemia: genomic
 structure and spectrum of mutations involving sterolin-1 and
 sterolin-2, encoded by ABCG5 and ABCG8, respectively

JOURNAL Am. J. Hum. Genet. 69 (2), 278-290 (2001)

MEDLINE 21344600

PUBMED 11452359

REFERENCE 2 (bases 1 to 3674)

AUTHORS Lu,K., Lee,M.-H. and Patel,S.B.

TITLE Direct Submission

JOURNAL Submitted (29-NOV-2000) Division of Endocrinology, Diabetes and
 Medical Genetics, Medical University of South Carolina, 114 Doughty

Street, STB541, Charleston, SC 29403, USA

FEATURES

source 1. .3674
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /tissue_type="liver"

gene 1. .3674
 /gene="Abcg8"

CDS 102. .2123
 /gene="Abcg8"
 /note="ABCG8"
 /codon_start=1
 /product="sterolin-2"
 /protein_id="AAK84079.1"
 /db_xref="GI:15088542"
 /translation="MAEKTKEETQLWNGTVLQDASQGLQDSLFSSSEDNSLYFTYSQGSNTLEVRDLTYQVDIASQVPWFEQLAQFKIPWRSHSSQDSCELGIRNLSFKVRSQMLAIIGSSGCGRASLLDVITGRGHGGKMKSGQIWINQGPSTPQLVRKCVAVHRQHDQLLPNLTVRETAFIAQMRLPRTFSQAQRDKRVEDVIAELRLRQCANTRVGNTYVRGVSGGERRRVSVIGVQLLWNPGLILDEPTSGLDSTAHNLVTTLSRLAKGNRLVLISLHQPRSDIFRLFDLVLLMTSGTPIYLGAQQMVQYFTSIGHPCPRYSNPADFVYDLTSDRRSKEREVATVEKAQSLAALFLEKVQGFDDFLWKAELAKELNTSTHTVSLTLTQDTDCGTAVELPGMIEQFSTLIRROIENDFRDLPTLLIHGSEACLSLIIGFLYYGHGAKQLSFMDTAALEFMIGALIPFNVILDVVSCKHSERSMLYYELEDGLYTAGPYFFAKILGELPEHCAYVIIYAMPIYWLTNLRPVPELFLHFLLVVLVVFCCRTMALAASAMLPTFHMSFFCNALYNSFYLTAGFMINLDNLWIWPAWISKLSFLRWCFSGLMQIQFNHLYTTQIGNFTFSILGDTMISAMDLSHPLYAIYLVIGISYGFLLYLLSLKLIKQKSIQDW"

ORIGIN

Query Match 100.0%; Score 102; DB 10; Length 3674;
 Best Local Similarity 100.0%; Pred. No. 1.2e-23;
 Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 269 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 210

Qy 61 ACTGTTGTCACTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 102
 |||||||||||||||||||||||||||||||||||||
 Db 209 ACTGTTGTCACTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 168

RESULT 5

AX685737

LOCUS AX685737 6043 bp DNA linear PAT 29-MAR-2003

DEFINITION Sequence 9 from Patent WO02081691.

ACCESSION AX685737

VERSION AX685737.1 GI:29371746

KEYWORDS .

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Hobbs,H.H., Shan,B., Barnes,R. and Tian,H.

TITLE Abcg5 and abcg8: compositions and methods of use
 JOURNAL Patent: WO 02081691-A 9 17-OCT-2002;
 Tularik Inc. (US) ; BOARD OF REGENTS UNIVERSITY OF TEXAS SYSTEM
 (US)

FEATURES Location/Qualifiers
 source 1. .6043
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="ABCG8 exon 2 (reverse strand) through ABCG5 exon 2
 (forward strand)"

ORIGIN

Query Match 100.0%; Score 102; DB 6; Length 6043;
 Best Local Similarity 100.0%; Pred. No. 1.2e-23;
 Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 3 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 62

Qy 61 ACTGTTGTCACTTTCCGAGGAGAAACAAGCTGTCCTGGAGGCC 102
 ||||||||||||||||||||||||||||||||||||
 Db 63 ACTGTTGTCACTTTCCGAGGAGAAACAAGCTGTCCTGGAGGCC 104

RESULT 6
 AY196215/c

LOCUS AY196215 2285 bp mRNA linear ROD 01-JUN-2003
 DEFINITION Mus musculus strain I/LnJ ATP-binding cassette sub-family G member
 8 (Abcg8) mRNA, complete cds.
 ACCESSION AY196215
 VERSION AY196215.1 GI:31322259
 KEYWORDS .
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2285)
 AUTHORS Wittenburg,H., Lyons,M.A., Li,R., Churchill,G.A., Carey,M.C. and
 Paigen,B.
 TITLE Primary Roles of FXR and ABCG5/ABCG8 in Cholesterol Gallstone
 Susceptibility: Evidence from a Cross of PERA/Ei and I/Ln Inbred
 Mice
 JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2285)
 AUTHORS Lyons,M.A., Wittenburg,H., Walsh,K.A., Carey,M.C. and Paigen,B.
 TITLE Direct Submission
 JOURNAL Submitted (12-DEC-2002) The Jackson Laboratory, 600 Main Street,
 Bar Harbor, ME 04609, USA

FEATURES Location/Qualifiers
 source 1. .2285
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="I/LnJ"
 /db_xref="taxon:10090"
 /chromosome="17"

```

/ map="55 cM"
/ sex="male"
/ tissue_type="liver"
gene      1. .2285
/ gene="Abcg8"
CDS       102. .2120
/ gene="Abcg8"
/ note="ATP-dependent canalicular cholesterol transporter;
white subfamily"
/ codon_start=1
/ product="ATP-binding cassette sub-family G member 8"
/ protein_id="AA045095.1"
/ db_xref="GI:31322260"
/ translation="MAEKTKEETQLWNGTVLQDASGLQDSLFSSES DNSLYFTYSGQS
NTLEVRDLTYQVDIASQVPWFQQLAQFKIPWRSHSSQDSCELGIRNLSFKVRSGQMLA
IIGSSGCGRASLLDVITGRGHGKMKSGQIWINGQPSTPQLVRKCAHVQRHDQLLPN
LTVRETLAFIAQMRLPRTFSQAQRDKRVEDVIAELRLRQCANTRVGNTYVRGVSGGER
RRVSIGVQLLWNP GILILDEPTSGLD SFTAHLNLTLSRLAKGNRLVLISLHQPRSDI
FRLFDLVLLMTSGTPIYLGAAQ QMVQYFTSIGHPCPRYSNPADFYVDLTSIDRRSKER
EVATVEKAQSLAALFLEKVQGFDDFLWKAEAKELNTSTHTVSLTLTQDTCGTAAELP
GMIEQFSTLIRROI S NDFRDLPTLLIHGSEACLSLIIGFLYYGHGAKQLSFMDTAAL
LFMIGALIPFNVILDV VSKCHSERSMLYYELEDGLYTAGPYFFAKILGELPEHCAYVI
IYAMPIYWLTNLRPVPELFLHLLLVLVVFCRTMALAASAMLP FHMSSFFCNALY
NSFYLTAGFMINLDNLWIVPAWISKLSFLRWCFSGLMQIQFNHLYTTQIGNFTFSIL
GDTMISAMD LNSHPLYAIYLIVIGISYGF LFLYYLSLKLKQKSIQDW"

```

ORIGIN

```

Query Match          98.4%;  Score 100.4;  DB 10;  Length 2285;
Best Local Similarity 99.0%;  Pred. No. 4e-23;
Matches 101;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

```

```

Qy      1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAG 60
        |||
Db      266 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCGCTGTAGGTGAAGTACAG 207

Qy      61 ACTGTTGTCACTTTCCGAGGAGACAAGCTGTCCTGGAGGCC 102
        |||
Db      206 ACTGTTGTCACTTTCCGAGGAGACAAGCTGTCCTGGAGGCC 165

```

RESULT 7

AF351785/c

LOCUS AF351785 4829 bp mRNA linear ROD 26-AUG-2002

DEFINITION Rattus norvegicus sterolin-2 (Abcg8) mRNA, complete cds.

ACCESSION AF351785

VERSION AF351785.2 GI:22477145

KEYWORDS .

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 4829)

AUTHORS Lu,K., Lee,M.H., Hazard,S., Brooks-Wilson,A., Hidaka,H., Kojima,H.,
Ose,L., Stalenhoef,A.F., Mietinnen,T., Bjorkhem,I., Bruckert,E.,
Pandya,A., Brewer,H.B. Jr., Salen,G., Dean,M., Srivastava,A. and
Patel,S.B.

TITLE Two genes that map to the STSL locus cause sitosterolemia: genomic structure and spectrum of mutations involving sterolin-1 and sterolin-2, encoded by ABCG5 and ABCG8, respectively
 JOURNAL Am. J. Hum. Genet. 69 (2), 278-290 (2001)
 MEDLINE 21344600
 PUBMED 11452359
 REFERENCE 2 (bases 1 to 4829)
 AUTHORS Lu,K., Yu,H., Lee,M. and Patel,S.B.
 TITLE Molecular cloning, genomic structure, and characterization of novel mouse head-to-head tandem ABC transporters
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 4829)
 AUTHORS Lu,K., Lee,M. and Patel,S.B.
 TITLE Direct Submission
 JOURNAL Submitted (21-FEB-2001) Division of Endocrinology, Diabetes and Medical Genetics, Medical University of South Carolina, 114 Doughty St, STB 541, Charleston, SC 29407, USA
 REFERENCE 4 (bases 1 to 4829)
 AUTHORS Lu,K., Yu,H., Lee,M. and Patel,S.B.
 TITLE Direct Submission
 JOURNAL Submitted (26-AUG-2002) Division of Endocrinology, Diabetes and Medical Genetics, Medical University of South Carolina, 114 Doughty St, STB 541, Charleston, SC 29403, USA
 REMARK Sequence update by submitter
 COMMENT On Aug 26, 2002 this sequence version replaced gi:15148516.
 FEATURES
 source Location/Qualifiers
 1. .4829
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 gene 1. .4829
 /gene="Abcg8"
 CDS 111. .2129
 /gene="Abcg8"
 /codon_start=1
 /product="sterolin-2"
 /protein_id="AAK84831.2"
 /db_xref="GI:22477146"
 /translation="MAEKTKEETQLWNGTVLQDASSLQDSVFSSSESDNSLYFTYSGQS
 NTLEVRDLTYQVDMASQVPWFEEQLAQFKLPWRSRGSQDSWDLGIRNLSFKVRSQGMLA
 IIGSAGCGRATLLDVITGRDHGGKMKSGQIWINQGPSTPQLIQKCVAHVRQQDQLLPN
 LTVRETLTFIAQMRLPKTFSQAQRDKRVEDVIAELRLRQCANTRVGNTYVRGVSGGER
 RRVSIGVQLLWNPGLILDEPTSGLDSTAHNLVRTLSRLAKGNRLVLISLHQPRSDI
 FRLFDLVLLMTSGTPIYLGVAQHMVQYFTSIGYPCPRYSNPADFYVDLTSIDRRSKEQ
 EVATMEKARLLAALFLEKVGQFDDFLWKAEAKSLDTGTYAVSQTLTQDTNCGTAAELP
 GMIQQFTTLIRROIISNDFRDLPTLFIHGAEACLSLIIGFLYYGHADKPLSFMDMAAL
 LFMIGALIPFNVILDVVSCKHSERSLLYYELEDGLYTAGPYFFAKVLGELPEHCAYVI
 IYGMPIYWLTNLRPGPELFLHFMLLWLWVFCRTMALAASAMLPFTHMSSFCCNALY
 NSFYLTAGFMINLNNLWIVPAWISKMSFLRWCFSGLMQIQFNHGYTTQIGNLTFSPV
 GDAMVTAMDLSHPLYAIYLIVIGISCGFLSLYYLSLKFIRKQKSIQDW"

ORIGIN

Query Match 92.7%; Score 94.6; DB 10; Length 4829;
 Best Local Similarity 96.0%; Pred. No. 3.8e-21;
 Matches 97; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
 |||
 Db 275 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTAGAG 216

Qy 61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGC 101
 |||
 Db 215 GCTGTTGTCACTTTCCAGAGGAGAACACGCTGTCCTGGAGGC 175

RESULT 8

AY145899

LOCUS AY145899 40929 bp DNA linear ROD 12-NOV-2002

DEFINITION Rattus norvegicus sterolin 2 (Abcg8) and sterolin 1 (Abcg5) genes, complete cds.

ACCESSION AY145899

VERSION AY145899.1 GI:24935208

KEYWORDS .

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 40929)

AUTHORS Yu,H., Lu,K., Lee,M., Pandit,B. and Patel,s.B.

TITLE The rat Abcg5 and Abcg8: characterization, chromosomal assignment and genetic variation in sitosterolemic rats

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 40929)

AUTHORS Yu,H., Lu,K., Lee,M., Pandit,B. and Patel,s.B.

TITLE Direct Submission

JOURNAL Submitted (29-AUG-2002) Endocrinology, Diabetes and Medical Genetics, Medical University of South Carolina, 114 Doughty Street, STR 541, Charleston, SC 29403, USA

FEATURES Location/Qualifiers

source 1. .40929
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 gene complement(<4136. .>20831)
 /gene="Abcg8"

mRNA complement(join(<4136. .4273,4361. .4488,5693. .5960, 6513. .6589,6754. .6953,8189. .8269,8350. .8512,10772. .11041,

11129. .11261,11647. .11885,15513. .15669,17473. .17574, 20769. .>20831))
 /gene="Abcg8"

CDS complement(join(4136. .4273,4361. .4488,5693. .5960, 6513. .6589,6754. .6953,8189. .8269,8350. .8512,10772. .11041,

11129. .11261,11647. .11885,15513. .15669,17473. .17574, 20769. .20831))
 /gene="Abcg8"
 /note="ATP-binding cassette sub-family G (WHITE) member 8"
 /codon_start=1
 /product="sterolin 2"

```

/protein_id="AAN64276.1"
/db_xref="GI:24935210"
/translation="MAEKTKEETQLWNGTVLQDASSLQDSVFSSES DNSLYFTYSGQS
NTLEVRDLTYQVDMASQVPWFELAQFKLPWRSRGSQDSWDLGIRNLSFKVRSQGMLA
IIGSAGCGRATLLDVITGRDHGGKMKSGQIWINQGPSTPQLIQKCVAVHRQQDQLLPN
LTVRETTLTFIAQMRLPKTFESQAQRDKRVEDVIAELRLRQCANTRVGNTYVRGVSGGER
RRVSIGVQLLWNPGLILDEPTSGLDSTAHNLVRTLSRLAKGNRLVLISLHQPRSDI
FRLFDLVLLMTSGTPPIYLGVAQHMVQYFTSIGYPCPRYSNPADFYVDLTSIDRRSKEQ
EVATMEKARLLAALFLEKVQGFDDFLWKAESA SLDGTGYAVSQTLTQDTNCGTAAELP
GMIQQFTTLIRRQISNDFRDLPTLFIHGAELMSLIIGFLYYGHADKPLSFMDMAAL
LFMIGALIPFNVILDVSKCHSERSLLYYELEDGLYTAGPYFFAKVLGELPEHCAYVI
IYGMPIYWLTNLRPGPELFLHFMLLWLVLVVFCCRTMALAASAMLPFHMSSFCCNALY
NSFYLTAGFMINLNNLWIVPAWISKMSFLRWCFSGLMQIQFNHGIYTTQIGNLTFSVP
GDAMVTAMDLSNHPYAIYLIVIGISCGFLSLYYLSLKFIKQKSIQDW"
gene
<21211. .>40564
/gene="Abcg5"
mRNA
join(<21211. .21356,21968. .22089,24726. .24862,24949.
.25047,
27388. .27520,28838. .28977,29879. .30008,30715. .30928,
31032. .31237,32869. .33007,35821. .36006,38553. .38665,
40371. .>40564)
/gene="Abcg5"
/product="sterolin 1"
CDS
join(21211. .21356,21968. .22089,24726. .24862,24949.
.25047,
27388. .27520,28838. .28977,29879. .30008,30715. .30928,
31032. .31237,32869. .33007,35821. .36006,38553. .38665,
40371. .40564)
/gene="Abcg5"
/note="ATP-binding cassette sub-family G (WHITE) member 5"
/codon_start=1
/product="sterolin 1"
/protein_id="AAN64275.1"
/db_xref="GI:24935209"
/translation="MSELPFLSPEGARGPHNNRGSQSSLEEGSVTGSEARHSLGVLNV
SFSVSNRVGPWWNIKSCQQKWDRIKLDVSLYIESGQTMCILGSSGSGKTTLLDAISG
RLRRTGTLEGEVFVNGCELRRDQFQDCVSYLLQSDVFLSSLTVRETLLRYTAMLALRSS
SADFYDKKVEAVLTELSSLHVADQMIGNYNFGGISSGERRRVSIAAQLQDPKVMMLD
EPTTGLDGMTANHIVLLVLELARRNRIVVTIHHQPRSELFHHFDKIAILTYGELVFCC
TPEEMLGFFNCCGYPCPEHSNPFDFYMDLTSVDTQSREREIETYKRVQMLESQFRQSD
ICHKILENIERTHLKTLPMVPEFTKNPPGMFCKLGVLRLRRVTRNLMRNKQVVMRLV
QNLIMGLFLIFYLLRVQNNMLKGAVQDRVGLLYQLVGATPYTGMLNAVNLFPMRLAVS
DQESQDGLYQKWQMLLAYVLHALPFSIVATVIFSSVCYWTGLYPEVARFGYFSAALL
APHLIGEFLTLVLLGMVQNPINVSIVALLSISGLLIGSGFIRNIEEMPIPLKILGYF
TFQKYCCAILVNEFYGLNFTCGGSNTSVPNPNMCSMTQGIQFIEKTCPGATSRFTTN
FLILYSFIPTLVILGMVVFKVRDYLISR"

```

ORIGIN

```

Query Match          92.7%;  Score 94.6;  DB 10;  Length 40929;
Best Local Similarity 96.0%;  Pred. No. 3.8e-21;
Matches 97;  Conservative 0;  Mismatches 4;  Indels 0;  Gaps 0;

Qy      1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      17473 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTAGAG 17532

Qy      61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGC 101

```

|||||
Db 17533 GCTGTTGTCACCTTTCAGAGGAGAACACGCTGTCTGGAGGC 17573

RESULT 9

AC120701/c

LOCUS AC120701 237445 bp DNA linear HTG 21-SEP-2002

DEFINITION Rattus norvegicus clone CH230-65H6, *** SEQUENCING IN PROGRESS ***,
4 unordered pieces.

ACCESSION AC120701

VERSION AC120701.4 GI:23265381

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 237445)

AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hagues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,

Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
 Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
 Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
 Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Weinstock,G. and Gibbs,R.A.

TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 237445)
 AUTHORS Worley,K.C.
 TITLE Direct Submission
 JOURNAL Submitted (09-MAY-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 REFERENCE 3 (bases 1 to 237445)
 AUTHORS Rat Genome Sequencing Consortium.
 TITLE Direct Submission
 JOURNAL Submitted (21-SEP-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 COMMENT On Sep 21, 2002 this sequence version replaced gi:21908396.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). As a result, the
 sequence may extend beyond the ends of the clone and there may be
 contigs that consist entirely of whole genome shotgun sequence
 reads. Both end sequences and whole genome shotgun sequence only
 contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GXQV
 Center clone name: CH230-65H6

----- Summary Statistics

Assembly program: Phrap; version 0.990329
 Consensus quality: 209781 bases at least Q40
 Consensus quality: 213033 bases at least Q30
 Consensus quality: 214997 bases at least Q20
 Estimated insert size: 233017; sum-of-contigs estimation
 Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will

* be preserved.
 * 1 233866: contig of 233866 bp in length
 * 233867 233966: gap of unknown length
 * 233967 235011: contig of 1045 bp in length
 * 235012 235111: gap of unknown length
 * 235112 236137: contig of 1026 bp in length
 * 236138 236237: gap of unknown length
 * 236238 237445: contig of 1208 bp in length.

FEATURES Location/Qualifiers
 source 1. .237445
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-65H6"
 misc_feature 1. .1326
 /note="wgs_end_extension
 clone_end:T7"
 misc_feature 8065. .8944
 /note="clone_boundary
 clone_end:T7
 site:EcoRI
 end_sequence:BH350813"
 misc_feature complement(232953. .233569)
 /note="clone_boundary
 clone_end:Sp6
 site:EcoRI
 end_sequence:BH350815"

ORIGIN

Query Match 92.7%; Score 94.6; DB 2; Length 237445;
 Best Local Similarity 96.0%; Pred. No. 3.9e-21;
 Matches 97; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAG 60
 |||

Db 141137 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTAGAG
 141078

Qy 61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGC 101
 |||

Db 141077 GCTGTTGTCACTTTCCGAGGAGAACACGCTGTCCTGGAGGC 141037

RESULT 10

AC112747

LOCUS AC112747 312858 bp DNA linear HTG 08-OCT-2002

DEFINITION Rattus norvegicus clone CH230-359E1, *** SEQUENCING IN PROGRESS
 ***, 8 unordered pieces.

ACCESSION AC112747

VERSION AC112747.3 GI:23270105

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 312858)

AUTHORS	<p> Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A. </p>
TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	2. (bases 1 to 312858)
AUTHORS	Worley,K.C.
TITLE	Direct Submission
JOURNAL	Submitted (24-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 312858)

AUTHORS Rat Genome Sequencing Consortium.

TITLE Direct Submission

JOURNAL Submitted (08-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT On Sep 23, 2002 this sequence version replaced gi:21738477. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GRAX

Center clone name: CH230-359E1

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 241372 bases at least Q40

Consensus quality: 245333 bases at least Q30

Consensus quality: 248022 bases at least Q20

Estimated insert size: 276767; sum-of-contigs estimation

Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)

* NOTE: This sequence may represent more than one clone.

* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 155105: contig of 155105 bp in length

* 155106 155205: gap of unknown length

* 155206 221765: contig of 66560 bp in length

* 221766 221865: gap of unknown length

* 221866 290378: contig of 68513 bp in length

* 290379 290478: gap of unknown length

* 290479 293724: contig of 3246 bp in length

* 293725 293824: gap of unknown length

* 293825 305790: contig of 11966 bp in length

* 305791 305890: gap of unknown length

* 305891 307341: contig of 1451 bp in length

* 307342 307441: gap of unknown length
 * 307442 309768: contig of 2327 bp in length
 * 309769 309868: gap of unknown length
 * 309869 312858: contig of 2990 bp in length.

FEATURES Location/Qualifiers
 source 1. .312858
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-359E1"
 misc_feature 159838. .161520
 /note="wgs_contig"
 misc_feature 166727. .168287
 /note="wgs_contig"
 misc_feature 190162. .191648
 /note="wgs_contig"
 misc_feature 234118. .235251
 /note="wgs_contig"
 misc_feature 290479. .292119
 /note="wgs_contig"

ORIGIN

Query Match 92.7%; Score 94.6; DB 2; Length 312858;
 Best Local Similarity 96.0%; Pred. No. 3.9e-21;
 Matches 97; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
 |||||
 Db 88051 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTAGAG 88110
 Qy 61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGC 101
 |||||
 Db 88111 GCTGTTGTCACTTTCAGAGGAGAACACGCTGTCCTGGAGGC 88151

RESULT 11 AF320294/c

LOCUS AF320294 2022 bp mRNA linear PRI 13-DEC-2000
 DEFINITION Homo sapiens ABCG8 (ABCG8) mRNA, complete cds.
 ACCESSION AF320294
 VERSION AF320294.1 GI:11692801
 KEYWORDS .
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2022)
 AUTHORS Berge,K.E., Tian,H., Graf,G.A., Yu,L., Grishin,N.V., Schultz,J.,
 Kwiterovich,P., Shan,B., Barnes,R. and Hobbs,H.H.
 TITLE Accumulation of Dietary Cholesterol in Sitosterolemia Caused by
 Mutations in Adjacent ABC Transporters
 JOURNAL Science (2001) In press
 REFERENCE 2 (bases 1 to 2022)
 AUTHORS Berge,K.E., Tian,H., Graf,G.A., Yu,L., Grishin,N.V., Schultz,J.,
 Kwiterovich,P., Shan,B., Barnes,R. and Hobbs,H.H.
 TITLE Direct Submission
 JOURNAL Submitted (09-NOV-2000) Molecular Genetics, University of Texas,

Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd.,
Dallas, TX 75390-9046, USA

FEATURES Location/Qualifiers
 source 1. .2022
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 gene 1. .2022
 /gene="ABCG8"
 CDS 1. .2022
 /gene="ABCG8"
 /note="ATP-binding cassette, subfamily G, member 8"
 /codon_start=1
 /product="ABCG8"
 /protein_id="AAG40004.1"
 /db_xref="GI:11692802"
 /translation="MAGKAAEERGLPKGATPQDTSGLQDRLFSSSEDNSLYFTYSGQP
NTLEVRDLNYQVDLASQVPWFQQLAQFKMPWTSPSCQNSCELGIQNLSFKVRSQGMLA
IIGSSGCGRASLLDVITGRGHGGKIKSGQIWINQPSQPQLVRKCVAHVRQHNQLLPN
LTVRETLAFIAQMRLPRTFSQAQRDKRVEDVIAELRLRQCADTRVGNMYVRGLSGGER
RRVSIGVQLLWNPGLILILDEPTSGLDSFTAHLNVLKLSRLAKGNRLVLISLHQPRSDI
FRLFDLVLLMTSGTPIYLGAAQHMVQYFTAIGYPCPRYSNPADFYVDLTSIDRRSREQ
ELATREKAQSLAALFLEKVRDLDDFLWKAETKDLEDTCVESSVTPLDNCPLSPPTKM
PGAVQQFTTLIRRQISNDFRDLPTLLIHGAEACLSMTIGFLYFGHGSIQLSFMDTAA
LLFMIGALIPFNVILDVISKCYSERAMLYYELEDGLYTTGPYFFAKILGELPEHCAYI
IIYGMPTYWLANLRPGLQPFLHFLVWLVVFCCRIMALAAAALLPTFHMAFFSNAL
YNSFYLAGGFMINLSSLWTVPWISKVSFLRWCFEGLMKIQFSRRTYKMPLGNLTIIV
SGDKILSVMELDSYPLYAIYLIVIGLSSGGFMVLYYVSLRFIKQKPSQDW"

ORIGIN

Query Match 85.9%; Score 87.6; DB 9; Length 2022;
Best Local Similarity 91.2%; Pred. No. 8.9e-19;
Matches 93; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
 ||||| |||| |||||||||||||| ||||| || |||||||||||||||||
Db 165 CTGGTAGTTGAGGTCTCTGACCTCCAGGGTGTGGGCTGGCCACTGTAGGTGAAGTACAG 106

Qy 61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 102
 |||||||||||| ||||||||| || |||||||||
Db 105 GCTGTTGTCACTTTTCAGAGGAGAACAATCTATCCTGGAGGCC 64

RESULT 12 AX685735/c

LOCUS AX685735 2669 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 7 from Patent WO02081691.
ACCESSION AX685735
VERSION AX685735.1 GI:29371744
KEYWORDS .
SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
 AUTHORS Hobbs,H.H., Shan,B., Barnes,R. and Tian,H.
 TITLE Abcg5 and abcg8: compositions and methods of use

JOURNAL Patent: WO 02081691-A 7 17-OCT-2002;
Tularik Inc. (US) ; BOARD OF REGENTS UNIVERSITY OF TEXAS SYSTEM
(US)

FEATURES Location/Qualifiers
source 1. .2669
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS 100. .2121
/note="unnamed protein product; human ABCG8 (hABCG8)"
/codon_start=1
/protein_id="CAD86573.1"
/db_xref="GI:29371745"
/db_xref="REMTREMBL:CAD86573"
/translation="MAGKAAEERGLPKGATPQDTSGLQDRLFSSSESDNSLYFTYSGQP
NTLEVRDLNYQVDLASQVPWFQLAQFKMPWTSPSCQNSCELGIQNLSFKVRSQQLA
IIGSSGCGRASLLDVTGRGHGGKIKSGQIWINQPPSPQLVRKCVAHVRQHNQLLPN
LTVRETLAFIAQMRLPRTFSQAQRDKRVEDVIAELRLRQCADTRVGNMYVRGLSGGER
RRVSIGVQLLWNPGLILDEPTSGLDSTAHNLVKTLRLAKGNRLVLISLHQPRSDI
FRLFDLVLMTSGTPIYLGAAQHMVQYFTAIGYPCPRYSNPADFYVDLTSIDRSREQ
ELATREKAQSLAALFLEKVRDLDDFLWKAETKDLDEDTCESSVTPLDTNCLPSPTKM
PGAVQQFTTLIRRQISNDFRDLPTLLIHGAACLMSMTIGFLYFGHGSIQLSFMDTAA
LLFMIGALIPFNVILDVISKYSERAMLYYELEDGLYTTGPYFFAKILGELPEHCAYI
IIYGMPTYWLANLRPGLQPFLLHFLLVLVVFCCRIMALAAAALLPTFHMASFFSNAL
YNSFYLAGGFMINLSSLWTVPAWISKVSFLRWCFEGLMKIQFSRRTYKMPLGNLTIAV
SGDKILSAMELDSYPLYAIYLIVIGLSGGFMVLYYVSLRFIKQKPSQDW"

ORIGIN

Query Match 85.9%; Score 87.6; DB 6; Length 2669;
Best Local Similarity 91.2%; Pred. No. 8.9e-19;
Matches 93; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
||||| ||| ||||||||| ||||| ||| |||||||||
Db 264 CTGGTAGTTGAGGTCTCTGACCTCCAGGGTGTGGGCTGGCCACTGTAGGTGAAGTACAG 205
Qy 61 ACTGTTGTCACTTTCAGAGGAGAACAAGCTGTCCTGGAGGCC 102
||||||| ||||||||| || |||||||||
Db 204 GCTGTTGTCACTTTCAGAGGAGAACAATCTATCCTGGAGGCC 163

RESULT 13 AC108476/c

LOCUS AC108476 139342 bp DNA linear PRI 16-APR-2002
DEFINITION Homo sapiens BAC clone RP11-1413K20 from 2, complete sequence.
ACCESSION AC108476
VERSION AC108476.5 GI:19807988
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 139342)
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792

PUBMED 9847074
REFERENCE 2 (bases 1 to 139342)
AUTHORS Harkins,C., Haakenson,W. and Doebber,A.
TITLE The sequence of Homo sapiens BAC clone RP11-1413K20
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 139342)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (27-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 139342)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 139342)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 6 (bases 1 to 139342)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Mar 29, 2002 this sequence version replaced gi:18767626.

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH1413K20
-----.

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-489K22, 2000 bp overlap. Actual end is at base position 139342 of RP11-1413K20.

The region between 132012 and 132017 is covered only by a pcr product of clone DNA.

FEATURES	Location/Qualifiers
source	1. .139342 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="2" /map="2" /clone="RP11-1413K20" /clone_lib="RPCI-11"
misc_feature	55. .655 /note="match to EST AA203458 (NID:g1799169) zx58b04.r1"
misc_feature	93. .286 /note="match to EST AV689089 (NID:g10290952)"
misc_feature	93. .286 /note="similar to Mus musculus EST AI597378 (NID:g4606426) vj29c06.y1"
misc_feature	93. .279 /note="match to EST AV660973 (NID:g9881987)"
misc_feature	318. .653 /note="match to EST R00405 (NID:g750141) ye71e05.r1"
misc_feature	372. .633 /note="similar to Homo sapiens EST T97887 (NID:g747232) ye58h05.r1"
misc_feature	706. .708 /note="match to EST R00405 (NID:g750141) ye71e05.r1"
misc_feature	706. .707 /note="similar to Homo sapiens EST T97887 (NID:g747232) ye58h05.r1"
repeat_region	847. .1139 /rpt_family="Alu"
misc_feature	1867. .2047 /note="match to EST T39945 (NID:g647612) ya13g04.r1"
repeat_region	2234. .2616 /rpt_family="L2"
misc_feature	2983. .3121 /note="match to EST AV689089 (NID:g10290952)"
misc_feature	2983. .3121 /note="similar to Mus musculus EST AI597378 (NID:g4606426) vj29c06.y1"
misc_feature	3044. .3121 /note="match to EST T86384 (NID:g714736) yd77b08.r1"

misc_feature 4099. .4304
 /note="match to EST T86384 (NID:g714736) yd77b08.r1"
 misc_feature 4099. .4283
 /note="match to EST AV689089 (NID:g10290952)"
 misc_feature 4401. .4618
 /note="similar to Mus musculus EST BF162656
 (NID:g11042879)"
 misc_feature 4405. .4454
 /note="match to EST T86384 (NID:g714736) yd77b08.r1"
 misc_feature 4724. .5110
 /note="similar to Homo sapiens EST AV656623
 (NID:g9877637)"
 misc_feature 5075. .5204
 /note="similar to Mus musculus EST BF162656
 (NID:g11042879)"
 repeat_region 5495. .5657
 /rpt_family="MIR"
 repeat_region 5673. .5767
 /rpt_family="MIR"
 repeat_region 5774. .5813
 /rpt_family="(TTG)n"
 repeat_region 5816. .5958
 /rpt_family="Alu"
 repeat_region 5976. .6091
 /rpt_family="MIR"
 repeat_region 6162. .6485
 /rpt_family="Alu"
 misc_feature 6351. .6373
 /note="match to EST AA228345 (NID:g1849916) nc39d04.s1"
 misc_feature 6352. .6364
 /note="match to EST AI431309 (NID:g4302284) ar55b01.x1"
 misc_feature 6352. .6364
 /note="match to EST AI469772 (NID:g4331862) tm20f11.x1"
 misc_feature 6353. .6367
 /note="match to EST AI241685 (NID:g3837082) qu70f06.x1"
 misc_feature 6568. .6707
 /note="similar to Mus musculus EST BF162656
 (NID:g11042879)"
 misc_feature 6649. .6707
 /note="similar to Mus musculus EST BB598373
 (NID:g16450340)"
 repeat_region 7229. .7528
 /rpt_family="Alu"
 misc_feature 7940. .8549
 /note="similar to EST BM725726 (NID:g19047059)"
 misc_feature 8169. .8305
 /note="similar to Mus musculus EST BF162656
 (NID:g11042879)"
 misc_feature 8169. .8301
 /note="similar to Mus musculus EST BB598373
 (NID:g16450340)"
 repeat_region 8500. .8529
 /rpt_family="AT_rich"
 repeat_region 8540. .8868
 /rpt_family="Alu"
 repeat_region 8870. .9180
 /rpt_family="Alu"

repeat_region 10493. .10636
 /rpt_family="MIR"
 repeat_region 11195. .11376
 /rpt_family="MER1_type"
 repeat_region 11377. .11658
 /rpt_family="Alu"
 repeat_region 11659. .11799
 /rpt_family="MER1_type"
 misc_feature 11955. .12053
 /note="similar to Mus musculus EST BB598373
 (NID:g16450340)"
 misc_feature 11994. .12053
 /note="similar to Mus musculus EST AA239884 (NID:g1863923)
 mx81d01.r1"
 repeat_region 12086. .12109

Query Match 85.9%; Score 87.6; DB 9; Length 139342;
 Best Local Similarity 91.2%; Pred. No. 9.2e-19;
 Matches 93; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
 ||||| ||| |||||||||||| ||||| || ||||||||||||||||
 Db 24794 CTGGTAGTTGAGGTCTCTGACCTCCAGGGTGTGGGCTGGCCACTGTAGGTGAAGTACAG 24735

 Qy 61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 102
 |||||||||| |||||||||| || |||||||||
 Db 24734 GCTGTTGTCACTTTCCAGAGGAGAACAATCTATCCTGGAGGCC 24693

RESULT 14 AC145533

LOCUS AC145533 159346 bp DNA linear HTG 19-JUL-2003
 DEFINITION Lemur catta clone LB2-138H20, WORKING DRAFT SEQUENCE, 5 unordered
 pieces.
 ACCESSION AC145533
 VERSION AC145533.1 GI:32996774
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Lemur catta (ring-tailed lemur)
 ORGANISM Lemur catta
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Strepsirhini; Lemuridae; Lemur.
 REFERENCE 1 (bases 1 to 159346)
 AUTHORS Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hosseini,R.,
 Peng,Z., Malinov,I. and Rubin,E.M.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 159346)
 AUTHORS Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hosseini,R.,
 Peng,Z., Malinov,I. and Rubin,E.M.
 TITLE Direct Submission
 JOURNAL Submitted (19-JUL-2003) Genome Sciences, Lawrence Berkeley National
 Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA
 COMMENT Draft Sequence Produced by Berkeley PGA
 Web site: <http://pga.lbl.gov>
 Center Code: PGABERK
 Center Project Name: L105-138H20
 Bac Clone Name: LB2-138H20

Additional information on comparative analysis and ordering are available at:
http://pga.lbl.gov/cgi-bin/search_cvcgd?type=n&value=ABCG5
Funding agent: Programs for Genomic Applications (NHLBI)
if library name is LB1 to LB4, please see website
for a description: <http://www-gsd.lbl.gov/cheng/BAC.html>
These libraries are available through the BACPAC Resources Center:
<http://www.chori.org/bacpac/libraryres.htm> as LBNL-1 to LBNL-4.

Summary Statistics:
Sequencing vector: Plasmid; pUC18
Chemistry: Dye-terminator Big Dye
Assembly program: Phrap version 0.990329.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 16021: contig of 16021 bp in length
* 16022 16121: gap of unknown length
* 16122 40145: contig of 24024 bp in length
* 40146 40245: gap of unknown length
* 40246 77537: contig of 37292 bp in length
* 77538 77637: gap of unknown length
* 77638 114811: contig of 37174 bp in length
* 114812 114911: gap of unknown length
* 114912 159346: contig of 44435 bp in length.

FEATURES Location/Qualifiers
source 1. .159346
/organism="Lemur catta"
/mol_type="genomic DNA"
/db_xref="taxon:9447"
/clone="LB2-138H20"

ORIGIN

Query Match 85.9%; Score 87.6; DB 2; Length 159346;
Best Local Similarity 91.2%; Pred. No. 9.3e-19;
Matches 93; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
||||| |||| |||||||||||| |||| | ||| ||||||||||||||||
Db 84994 CTGGTAGCTGAGGTCTCTGACCTCCAGGGTGTAGGCTGGCCACTGTAGGTGAAGTACAG 85053

Qy 61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 102
||||||||||| ||||||||||||| |||||||||
Db 85054 GCTGTTGTCACTTTTCAGAGGAGAACAAGCTATCCTGGAGGCC 85095

RESULT 15
AF324494/c
LOCUS AF324494 2679 bp mRNA linear PRI 07-AUG-2001
DEFINITION Homo sapiens sterolin-2 (ABCG8) mRNA, complete cds.
ACCESSION AF324494
VERSION AF324494.1 GI:15088539

KEYWORDS .

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2679)

AUTHORS Lu, K., Lee, M.H., Hazard, S., Brooks-Wilson, A., Hidaka, H., Kojima, H.,
Ose, L., Stalenhoef, A.F., Mietinnen, T., Bjorkhem, I., Bruckert, E.,
Pandya, A., Brewer, H.B. Jr., Salen, G., Dean, M., Srivastava, A. and
Patel, S.B.

TITLE Two genes that map to the STSL locus cause sitosterolemia: genomic
structure and spectrum of mutations involving sterolin-1 and
sterolin-2, encoded by ABCG5 and ABCG8, respectively

JOURNAL Am. J. Hum. Genet. 69 (2), 278-290 (2001)

MEDLINE 21344600

PUBMED 11452359

REFERENCE 2 (bases 1 to 2679)

AUTHORS Lu, K., Lee, M.-H. and Patel, S.B.

TITLE Direct Submission

JOURNAL Submitted (29-NOV-2000) Division of Endocrinology, Diabetes and
Medical Genetics, Medical University of South Carolina, 114 Doughty
Street, STB541, Charleston, SC 29403, USA

FEATURES Location/Qualifiers

source 1. .2679
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2p21; between D2S2294 and D2S2298"
/tissue_type="liver"

gene 1. .2679
/gene="ABCG8"

CDS 91. .2112
/gene="ABCG8"
/codon_start=1
/product="sterolin-2"
/protein_id="AAK84078.1"
/db_xref="GI:15088540"
/translation="MAGKAAEERGLPKGATPQDTSGLQDRLFSSES DNSLYFTYSGQP
NTLEVRDLNCQVDLASQVPWFEEQLAQFKMPWTS PSCQNSCELGIQNLSFKVRS GQMLA
IIGSSGCGRASLLDVITGRGHGGKIKSGQI WINGQPSSPQLVRKCVAHVRQHNQLLPN
LTVRETLAFIAQMRLPRTFSQAQRDKRVEDVIAELRLRQCADTRVGNMYVRGLSGGER
RRVSIQVQLLWNP GILILDEPTSGLD SFTAHLVKLTSLRLAKGNRLVLISLHQPRSDI
FRLFDLVLLMTSGTPIYLGAAQHMVQYFTAIGYPCPRYSNPADFYVDLTSIDRRSREQ
ELATREKAQSLAALFLEKVRDLDDFLWKAETKDLDEDTCEVSSVTPLDTNCLPSP TKM
PGAVQQFTTLIRRQISNDFRDLPTLLIHGA EACLSMTIGFLYFGHGS IQLSFMDTAA
LLFMIGALIPFNVILDVISKYSERAMLYYELEDGLYTTGPYFFAKILGELPEHCAYI
IIYGMPTYWLANLRPGLQPFLLHFLLVVLVVFCCRIMALAAAALLPTFHMASFFSNAL
YNSFYLAGGFMINLSSLWTVPAWISKVSFLRWC FEGLMKIQFSRRTYKMPLGNLTIAV
SGDKILSAMELDSYPLYAIYLIVIGLSGGFMVLYYVSLRFIKQKPSQDW"

ORIGIN

Query Match 84.3%; Score 86; DB 9; Length 2679;
Best Local Similarity 90.2%; Pred. No. 3.1e-18;
Matches 92; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60

```

      ||| || ||| ||||| ||||| ||| ||||| ||||| |||||
Db      255 CTGGCAGTTGAGGTCTCTGACCTCCAGGGTGTGGGGCTGGCCACTGTAGGTGAAGTACAG 196

Qy      61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 102
      ||||| ||||| ||||| ||| ||||| |||||
Db      195 GCTGTTGTCACTTTTCAGAGGAGAACAATCTATCCTGGAGGCC 154

```

RESULT 16

F351812S02/c

LOCUS F351812S02 4665 bp DNA linear PRI 10-AUG-2001

DEFINITION Homo sapiens sterolin-2 (ABCG8) gene, exon 2.

ACCESSION AF351813

VERSION AF351813.1 GI:15146432

KEYWORDS .

SEGMENT 2 of 13

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 4665)

AUTHORS Lu,K., Lee,M.H., Hazard,S., Brooks-Wilson,A., Hidaka,H., Kojima,H.,
Ose,L., Stalenhoef,A.F., Mietinnen,T., Bjorkhem,I., Bruckert,E.,
Pandya,A., Brewer,H.B. Jr., Salen,G., Dean,M., Srivastava,A. and
Patel,S.B.

TITLE Two genes that map to the STSL locus cause sitosterolemia: genomic
structure and spectrum of mutations involving sterolin-1 and
sterolin-2, encoded by ABCG5 and ABCG8, respectively

JOURNAL Am. J. Hum. Genet. 69 (2), 278-290 (2001)

MEDLINE 21344600

PUBMED 11452359

REFERENCE 2 (bases 1 to 4665)

AUTHORS Lu,K.

TITLE Direct Submission

JOURNAL Submitted (21-FEB-2001) Division of Endocrinology, Diabetes and
Medical Genetics, Medical University of South Carolina, 114 Doughty
St, STB 541, Charleston, SC 29403, USA

FEATURES Location/Qualifiers

```

source      1. .4665
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"
             /chromosome="2"
             /map="between D2S2294 and D2S2298"
             /clone="1081G2; 328I4"
             /cell_type="ES cell"
exon        3941..4042
             /gene="ABCG8"
             /number=2

```

ORIGIN

Query Match 84.3%; Score 86; DB 9; Length 4665;

Best Local Similarity 90.2%; Pred. No. 3.1e-18;

Matches 92; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

Qy      1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
      ||| || ||| ||||| ||||| ||| ||||| ||||| |||||

```

Db 4042 CTGGCAGTTGAGGTCTCTGACCTCCAGGGTGTGGGCTGGCCACTGTAGGTGAAGTACAG 3983

Qy 61 ACTGTTGTCACTTTCCGAGGAGAACAAAGCTGTCCTGGAGGCC 102

|||||

Db 3982 GCTGTTGTCACTTTCCAGAGGAGAACAAATCTATCCTGGAGGCC 3941

MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
 Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission
 JOURNAL Submitted (24-AUG-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE 4 (bases 1 to 127066)
 AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
 Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
 Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
 Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
 Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission
 JOURNAL Submitted (11-DEC-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 COMMENT On Dec 11, 2001 this sequence version replaced gi:15284200.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L11578
 Center clone name: 2367_F_13

FEATURES Location/Qualifiers
 source 1. .127066
 /organism="Homo sapiens"


```

/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="CTB-2367F13"
/clone_lib="CITB Human BAC"
repeat_region complement(8. .170)
/rpt_family="MER47A"
repeat_region 171. .468
/rpt_family="AluSx"
repeat_region complement(469. .516)
/rpt_family="MER47A"
repeat_region 988. .1049
/rpt_family="MIR"
repeat_region complement(1294. .1448)
/rpt_family="L1ME4A"
repeat_region complement(2662. .2954)
/rpt_family="AluSx"
repeat_region 4049. .4431
/rpt_family="L2"
unsure 5261. .5269
/note="<30 qual SNGL region"
unsure 7192. .7202
/note="<30 qual SNGL region"
repeat_region 7310. .7472
/rpt_family="MIR"
repeat_region 7488. .7582
/rpt_family="MIR"
repeat_region 7589. .7628
/rpt_family="(TTG)n"
repeat_region complement(7631. .7781)
/rpt_family="AluSg/x"
repeat_region 7791. .7922
/rpt_family="MIR"
repeat_region complement(7977. .8300)
/rpt_family="AluJb"
repeat_region 9044. .9343
/rpt_family="AluSq"
repeat_region 10315. .10344
/rpt_family="AT_rich"
repeat_region 10355. .10681
/rpt_family="AluJo"
repeat_region 10683. .10993
/rpt_family="AluSx"
repeat_region complement(12221. .12282)
/rpt_family="MIR3"
repeat_region complement(12306. .12449)
/rpt_family="MIR"
repeat_region complement(13008. .13189)
/rpt_family="MER33"
repeat_region complement(13190. .13471)
/rpt_family="AluJo"
repeat_region complement(13472. .13612)
/rpt_family="MER33"
repeat_region 13899. .13922
/rpt_family="GC_rich"
repeat_region complement(14184. .14250)

```

```

repeat_region      /rpt_family="L2"
14552. .14630
repeat_region      /rpt_family="MER5A"
complement(14809. .15100)
repeat_region      /rpt_family="AluSx"
complement(15363. .15679)
repeat_region      /rpt_family="AluY"
complement(15681. .15979)
repeat_region      /rpt_family="AluSx"
16292. .16388
repeat_region      /rpt_family="L2"
16392. .16508
repeat_region      /rpt_family="MLT1F"
complement(16538. .16616)
repeat_region      /rpt_family="LTR37B"
16618. .16687
repeat_region      /rpt_family="Alu"
complement(16988. .17104)
repeat_region      /rpt_family="L2"
17540. .17895
repeat_region      /rpt_family="MLT1A1"
complement(17911. .18209)
repeat_region      /rpt_family="AluSq"
18487. .18680
repeat_region      /rpt_family="LTR16A1"
18802. .19026
repeat_region      /rpt_family="AluJo"
complement(19092. .19390)
repeat_region      /rpt_family="AluJo"
complement(21369. .21675)
repeat_region      /rpt_family="AluSx"
complement(22474. .22763)
repeat_region      /rpt_family="MER115"
complement(22843. .22942)
repeat_region      /rpt_family="MER115"
23239. .23311
repeat_region      /rpt_family="L3"
complement(23968. .24265)

```

```

Query Match          84.3%;  Score 86;  DB 9;  Length 127066;
Best Local Similarity 90.2%;  Pred. No. 3.2e-18;
Matches 92;  Conservative 0;  Mismatches 10;  Indels 0;  Gaps 0;

```

```

Qy      1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAG 60
        |||| || |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      26562 CTGGCAGTTGAGGTCTCTGACCTCCAGGGTGTTGGGCTGGCCACTGTAGGTGAAGTACAG 26503

Qy      61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 102
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      26502 GCTGTTGTCACTTTCCAGAGGAGAACAATCTATCCTGGAGGCC 26461

```

```

RESULT 18
AC146787/c
LOCUS      AC146787          178016 bp    DNA    linear    HTG 03-OCT-2003
DEFINITION Aotus nancymae clone CH258-323A5, WORKING DRAFT SEQUENCE, 4
            ordered pieces.

```

ACCESSION AC146787
 VERSION AC146787.1 GI:37497135
 KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
 SOURCE Aotus nancymaae (Ma's night monkey)
 ORGANISM Aotus nancymaae
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.
 REFERENCE 1 (bases 1 to 178016)
 AUTHORS Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R.,
 Peng, Z., Malinov, I. and Rubin, E.M.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 178016)
 AUTHORS Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R.,
 Peng, Z., Malinov, I. and Rubin, E.M.
 TITLE Direct Submission
 JOURNAL Submitted (03-OCT-2003) Genome Sciences, Lawrence Berkeley National
 Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA
 COMMENT

Sequence Produced by Berkeley PGA
 Web site: <http://pga.lbl.gov>
 Center Code: PGABERK
 Center Project Name: W010
 Bac Clone Name: CH258-323A5

This sequence has been compared to sequences of other species
 using Vista (<http://www-gsd.lbl.gov/VISTA>). The results can be
 viewed at:
http://pga.lbl.gov/cgi-bin/search_cvcgd?type=n&value=ABCG5

The order-orientation of the draft sequence was accomplished by
 using:
 Avid (<http://baboon.math.berkeley.edu/mavid>),
 Lagan (<http://lagan.stanford.edu/>) and paired end information.

Funding agent: Programs for Genomic Applications (NHLBI)

Summary Statistics:

Sequencing vector: Plasmid; pUC18

Chemistry: Dye-terminator Big Dye

Assembly program: Phrap version 0.990329.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 4 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 32150: contig of 32150 bp in length

* 32151 32250: gap of unknown length

* 32251 56222: contig of 23972 bp in length

* 56223 56322: gap of unknown length

* 56323 173105: contig of 116783 bp in length

* 173106 173205: gap of unknown length

* 173206 178016: contig of 4811 bp in length.

This sequence has been compared to sequences of other species using Vista (<http://www-gsd.lbl.gov/VISTA>). The results can be viewed at:
http://pga.lbl.gov/cgi-bin/search_cvcgd?type=n&value=ABCG5

The order-orientation of the draft sequence was accomplished by using:

Avid (<http://baboon.math.berkeley.edu/mavid>),

Lagan (<http://lagan.stanford.edu/>) and paired end information.

Funding agent: Programs for Genomic Applications (NHLBI)

Summary Statistics:

Sequencing vector: Plasmid; pUC18

Chemistry: Dye-terminator Big Dye

Assembly program: Phrap version 0.990329.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 3 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 49109: contig of 49109 bp in length

* 49110 49209: gap of unknown length

* 49210 57420: contig of 8211 bp in length

* 57421 57520: gap of unknown length

* 57521 185045: contig of 127525 bp in length.

FEATURES

source

Location/Qualifiers

1. .185045

/organism="Callithrix jacchus"

/mol_type="genomic DNA"

/db_xref="taxon:9483"

/clone="CH259-274K20"

ORIGIN

Query Match 82.7%; Score 84.4; DB 2; Length 185045;

Best Local Similarity 89.2%; Pred. No. 1.1e-17;

Matches 91; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60

||| ||| |||| |||||||||||| || |||| ||| ||||||||||||||||

Db 121318 CTGATAGTTGAGGTCTCTGACCTCCAGGGTATTGGGCTGGCCACTGTAGGTGAAGTACAG

121259

Qy 61 ACTGTTGTCACTTTCAGAGGAGAACAAGCTGTCCTGGAGGCC 102

|||||||||||| |||||||||| || |||||||||

Db 121258 GCTGTTGTCACTTTCAGAGGAGAACAATCTATCCTGGAGGCC 121217

RESULT 20

AC146464

LOCUS AC146464 202533 bp DNA linear HTG 19-AUG-2003

DEFINITION Saimiri sciureus clone CH254-84A11, WORKING DRAFT SEQUENCE.

ACCESSION AC146464

VERSION AC146464.1 GI:33636782

KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.

SOURCE Saimiri sciureus (common squirrel monkey)

ORGANISM Saimiri sciureus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.

REFERENCE 1 (bases 1 to 202533)

AUTHORS Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R., Peng, Z., Malinov, I. and Rubin, E.M.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 202533)

AUTHORS Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R., Peng, Z., Malinov, I. and Rubin, E.M.

TITLE Direct Submission

JOURNAL Submitted (14-AUG-2003) Genome Sciences, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA

REFERENCE 3 (bases 1 to 202533)

AUTHORS Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R., Peng, Z., Malinov, I. and Rubin, E.M.

TITLE Direct Submission

JOURNAL Submitted (19-AUG-2003) Genome Sciences, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA

COMMENT

Sequence Produced by Berkeley PGA
 Web site: <http://pga.lbl.gov>
 Center Code: PGABERK
 Center Project Name: S030
 Bac Clone Name: CH254-84A11

This sequence has been compared to sequences of other species using Vista (<http://www-gsd.lbl.gov/VISTA>). The results can be viewed at:
http://pga.lbl.gov/cgi-bin/search_cvcgd?type=n&value=ABCG5

The order-orientation of the draft sequence was accomplished by using:
 Avid (<http://baboon.math.berkeley.edu/mavid>),
 Lagan (<http://lagan.stanford.edu/>) and paired end information.

Funding agent: Programs for Genomic Applications (NHLBI)

Summary Statistics:

Sequencing vector: Plasmid; pUC18
 Chemistry: Dye-terminator Big Dye
 Assembly program: Phrap version 0.990329.

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES

source

1 202533: contig of 202533 bp in length.
 Location/Qualifiers
 1. .202533
 /organism="Saimiri sciureus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9521"

/clone="CH254-84A11"

ORIGIN

Query Match 81.2%; Score 82.8; DB 2; Length 202533;
Best Local Similarity 88.2%; Pred. No. 4e-17;
Matches 90; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
||| ||| |||| ||| ||||||||| || ||| ||| |||||||||
Db 27346 CTGATAGTTGAGGTCTTTGACCTCCAGGGTATTGGGCTGGCCACTGTAGGTGAAGTACAG 27405
Qy 61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 102
||||||||||| ||||||||| || |||||||||
Db 27406 GCTGTTGTCACTTTCCGAGGAGAACAATCTATCCTGGAGGCC 27447

RESULT 21

AC146286/c

LOCUS AC146286 207760 bp DNA linear HTG 15-AUG-2003

DEFINITION Callicebus moloch clone LB5-414K16, WORKING DRAFT SEQUENCE, 2
ordered pieces.

ACCESSION AC146286

VERSION AC146286.2 GI:33667134

KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.

SOURCE Callicebus moloch (Dusky titi)

ORGANISM Callicebus moloch

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Callicebinae;
Callicebus.

REFERENCE 1 (bases 1 to 207760)

AUTHORS Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R.,
Peng, Z., Malinov, I. and Rubin, E.M.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 207760)

AUTHORS Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R.,
Peng, Z., Malinov, I. and Rubin, E.M.

TITLE Direct Submission

JOURNAL Submitted (02-AUG-2003) Genome Sciences, Lawrence Berkeley National
Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA

REFERENCE 3 (bases 1 to 207760)

AUTHORS Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R.,
Peng, Z., Malinov, I. and Rubin, E.M.

TITLE Direct Submission

JOURNAL Submitted (15-AUG-2003) Genome Sciences, Lawrence Berkeley National
Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA

COMMENT On Aug 15, 2003 this sequence version replaced gi:33413351.

Sequence Produced by Berkeley PGA

Web site: <http://pga.lbl.gov>

Center Code: PGABERK

Center Project Name: T039

Bac Clone Name: LB5-414K16

This sequence has been compared to sequences of other species
using Vista (<http://www-gsd.lbl.gov/VISTA>). The results can be
viewed at:

http://pga.lbl.gov/cgi-bin/search_cvcgd?type=n&value=ABCG5

The order-orientation of the draft sequence was accomplished by using:

Avid (<http://baboon.math.berkeley.edu/mavid>),

Lagan (<http://lagan.stanford.edu/>) and paired end information.

Funding agent: Programs for Genomic Applications (NHLBI)

If the Bac Library Name is LB1 to LB4, please see website for the description: <http://www-gsd.lbl.gov/cheng/BAC.html>
These libraries are available through the BACPAC Resources Center: <http://www.chori.org/bacpac/libraryres.htm> as LBNL-1 to LBNL-4.

Summary Statistics:

Sequencing vector: Plasmid; pUC18

Chemistry: Dye-terminator Big Dye

Assembly program: Phrap version 0.990329.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

* 1 74764: contig of 74764 bp in length

* 74765 74864: gap of unknown length

* 74865 207760: contig of 132896 bp in length.

FEATURES

source

Location/Qualifiers

1. .207760

/organism="Callicebus moloch"

/mol_type="genomic DNA"

/db_xref="taxon:9523"

/clone="LB5-414K16"

ORIGIN

Query Match 81.2%; Score 82.8; DB 2; Length 207760;

Best Local Similarity 88.2%; Pred. No. 4e-17;

Matches 90; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60

||| ||| |||| ||||||||| || || |||| ||| |||||||||||||||||

Db 145281 CTGATAGTTGAGGTCTCTGACCTCTAGGGTATTGGGCTGGCCACTGTAGGTGAAGTACAG
145222

Qy 61 ACTGTTGTCACTTCCGAGGAGACAAGCTGTCCTGGAGGCC 102

||||||||||| ||||||||| || |||||||||

Db 145221 GCTGTTGTCACTTTCAGAGGAGACAATCTATCCTGGAGGCC 145180

RESULT 22

AC146282/c

LOCUS AC146282 135280 bp DNA linear HTG 02-AUG-2003

DEFINITION Takifugu rubripes clone MRC-186C24, WORKING DRAFT SEQUENCE, 7
unordered pieces.

ACCESSION AC146282
 VERSION AC146282.1 GI:33413347
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Takifugu rubripes (Fugu rubripes)
 ORGANISM Takifugu rubripes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetradontoidea; Tetraodontidae; Takifugu.
 REFERENCE 1 (bases 1 to 135280)
 AUTHORS Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R.,
 Peng, Z., Malinov, I. and Rubin, E.M.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 135280)
 AUTHORS Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R.,
 Peng, Z., Malinov, I. and Rubin, E.M.
 TITLE Direct Submission
 JOURNAL Submitted (02-AUG-2003) Genome Sciences, Lawrence Berkeley National
 Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA
 COMMENT Draft Sequence Produced by Berkeley PGA
 Web site: <http://pga.lbl.gov>
 Center Code: PGABERK
 Center Project Name: F069-186C24
 Bac Clone Name: MRC-186C24

Additional information on comparative analysis and ordering are available at:

http://pga.lbl.gov/cgi-bin/search_cvcgd?type=n&value=
 Funding agent: Programs for Genomic Applications (NHLBI)
 Summary Statistics:

Sequencing vector: Plasmid; pUC18
 Chemistry: Dye-terminator Big Dye
 Assembly program: Phrap version 0.990329.

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 7 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 28849: contig of 28849 bp in length
 * 28850 28949: gap of unknown length
 * 28950 40654: contig of 11705 bp in length
 * 40655 40754: gap of unknown length
 * 40755 55789: contig of 15035 bp in length
 * 55790 55889: gap of unknown length
 * 55890 70983: contig of 15094 bp in length
 * 70984 71083: gap of unknown length
 * 71084 90702: contig of 19619 bp in length
 * 90703 90802: gap of unknown length
 * 90803 112817: contig of 22015 bp in length
 * 112818 112917: gap of unknown length
 * 112918 135280: contig of 22363 bp in length.

FEATURES Location/Qualifiers
 source 1..135280

/organism="Takifugu rubripes"
/mol_type="genomic DNA"
/db_xref="taxon:31033"
/clone="MRC-186C24"

ORIGIN

Query Match 51.4%; Score 52.4; DB 2; Length 135280;
Best Local Similarity 75.6%; Pred. No. 8.3e-07;
Matches 65; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
|| ||| |||| || ||||| || ||| || ||||| ||||| ||
Db 34807 CTCATAGTTGAGGTCGTTGACCTCCAGCTGGTTGCACCCTCCACTGTAGGTGAAGTAGAG 34748

Qy 61 ACTGTTGTCACTTTCCGAGGAGAACA 86
||| |||| ||| | |||||
Db 34747 GCTGCTGTCTTCTTCAGTGGAGAACA 34722

RESULT 23

AL928999

LOCUS AL928999 169570 bp DNA linear VRT 24-DEC-2002

DEFINITION Zebrafish DNA sequence from clone CH211-227C6, complete sequence.

ACCESSION AL928999

VERSION AL928999.4 GI:26788223

KEYWORDS HTG.

SOURCE Danio rerio (zebrafish)

ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 169570)

AUTHORS Heath, P.

TITLE Direct Submission

JOURNAL Submitted (21-DEC-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zface@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk

COMMENT On Dec 13, 2002 this sequence version replaced gi:25055310.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: zface@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare

occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www/Projects/D_rerio/fishmask.shtml
CH211-227C6 is from a CHORI-211 BAC library
VECTOR: pTARBAC2.1.

FEATURES
source Location/Qualifiers
1. .169570
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-227C6"
/clone_lib="CHORI-211"

ORIGIN

Query Match 35.1%; Score 35.8; DB 5; Length 169570;
Best Local Similarity 63.2%; Pred. No. 0.36;
Matches 55; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
|| ||| ||||| | ||| ||||| ||| || ||||| || || |||||
Db 74347 CTCATAGTTGAGATTGCGGACTTCCAGTTCATTGCGGCCTCCGCTGTAAGTAAATACAG 74406
Qy 61 ACTGTTGTCACTTTCGAGGAGAACAA 87
|||| |||| || | || | ||
Db 74407 ACTGCTGTCCTCCTCTGGAGATGAAAA 74433

RESULT 24

BX004832/c
LOCUS BX004832 190952 bp DNA linear VRT 25-NOV-2003
DEFINITION Zebrafish DNA sequence from clone CH211-89M19, complete sequence.
ACCESSION BX004832
VERSION BX004832.9 GI:38524388
KEYWORDS HTG.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 190952)
AUTHORS Harrison,E.
TITLE Direct Submission
JOURNAL Submitted (25-NOV-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT On Nov 25, 2003 this sequence version replaced gi:31335509.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>

Contact: zfish-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see

http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml CH211-89M19 is from a CHORI-211 BAC library

VECTOR: pTARBAC2.1

Clone-derived Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

FEATURES	Location/Qualifiers
source	1..190952 /organism="Danio rerio" /mol_type="genomic DNA" /db_xref="taxon:7955" /clone="CH211-89M19" /clone_lib="CHORI-211"

ORIGIN

Query Match	33.5%;	Score 34.2;	DB 5;	Length 190952;
Best Local Similarity	62.1%;	Pred. No. 1.3;		
Matches	54;	Conservative	0;	Mismatches 33; Indels 0; Gaps 0;

Qy	1	CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG	60
Db	92088	CTCATAGTTGAGATTGCGGACTTCCAGTTCATTGCGGCCTCCGCTGTAAGTAAAATACAA	92029
Qy	61	ACTGTTGTCACTTTCCGAGGAGAACAA	87
Db	92028	ACTGCTGTCCTCCTCTGGAGATGAAAA	92002

RESULT 25

BX571838

LOCUS	BX571838	226929 bp	DNA	linear	HTG 27-SEP-2003
-------	----------	-----------	-----	--------	-----------------

DEFINITION Danio rerio clone DKEY-205N7, WORKING DRAFT SEQUENCE, 14 unordered pieces.

ACCESSION BX571838

VERSION BX571838.3 GI:36796624

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLLTOP.

SOURCE Danio rerio (zebrafish)

ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 226929)

AUTHORS McLaren, S.

TITLE Direct Submission

JOURNAL Submitted (26-SEP-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

COMMENT On Sep 27, 2003 this sequence version replaced gi:33386624.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zK205N7
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 223662 bases at least Q40
Consensus quality: 224399 bases at least Q30
Consensus quality: 224947 bases at least Q20
Insert size: 225629; sum-of-contigs
Insert size: 196940; 4.8% error; agarose-fp
Quality coverage: 6.66x in Q20 bases; sum-of-contigs Quality
coverage: 7.66x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 10067: contig of 10067 bp in length
* 10068 10167: gap of 100 bp
* 10168 24021: contig of 13854 bp in length
* 24022 24121: gap of 100 bp
* 24122 28447: contig of 4326 bp in length
* 28448 28547: gap of 100 bp
* 28548 47699: contig of 19152 bp in length
* 47700 47799: gap of 100 bp
* 47800 68972: contig of 21173 bp in length
* 68973 69072: gap of 100 bp
* 69073 73919: contig of 4847 bp in length
* 73920 74019: gap of 100 bp
* 74020 106234: contig of 32215 bp in length
* 106235 106334: gap of 100 bp

```

* 106335 126675: contig of 20341 bp in length
* 126676 126775: gap of 100 bp
* 126776 145072: contig of 18297 bp in length
* 145073 145172: gap of 100 bp
* 145173 161897: contig of 16725 bp in length
* 161898 161997: gap of 100 bp
* 161998 198188: contig of 36191 bp in length
* 198189 198288: gap of 100 bp
* 198289 204891: contig of 6603 bp in length
* 204892 204991: gap of 100 bp
* 204992 210028: contig of 5037 bp in length
* 210029 210128: gap of 100 bp
* 210129 226929: contig of 16801 bp in length.

```

```

FEATURES                      Location/Qualifiers
    source                    1. .226929
                              /organism="Danio rerio"
                              /mol_type="genomic DNA"
                              /db_xref="taxon:7955"
                              /clone="DKEY-205N7"
                              /clone_lib="DanioKey"
    misc_feature              1. .10067
                              /note="assembly_fragment:01322
                              fragment_chain:1"
    misc_feature              10168. .24021
                              /note="assembly_fragment:01600
                              fragment_chain:1"
    misc_feature              24122. .28447
                              /note="assembly_fragment:01034
                              fragment_chain:1"
    misc_feature              28548. .47699
                              /note="assembly_fragment:01370
                              fragment_chain:1"
    misc_feature              47800. .68972
                              /note="assembly_fragment:01889
                              fragment_chain:1"
    misc_feature              69073. .73919
                              /note="assembly_fragment:00479
                              fragment_chain:1"
    misc_feature              74020. .106234
                              /note="assembly_fragment:00303
                              fragment_chain:1"
    misc_feature              106335. .126675
                              /note="assembly_fragment:01196
                              fragment_chain:1"
    misc_feature              126776. .145072
                              /note="assembly_fragment:00768
                              fragment_chain:1"
    misc_feature              145173. .161897
                              /note="assembly_fragment:00378
                              fragment_chain:1"
    misc_feature              161998. .198188
                              /note="assembly_fragment:00694
                              fragment_chain:1"
    misc_feature              198289. .204891
                              /note="assembly_fragment:01048
                              fragment_chain:1"
    misc_feature              204992. .210028

```

misc_feature /note="assembly_fragment:00601
fragment_chain:1"
210129.226929
/note="assembly_fragment:01542.0"

ORIGIN

Query Match 33.5%; Score 34.2; DB 2; Length 226929;
Best Local Similarity 62.1%; Pred. No. 1.3;
Matches 54; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
|| ||| ||||| | ||| ||||| ||| || ||||| || || |||||
Db 220615 CTCATAGTTGAGATTGCGGACTTCCAGTTCATTGCGGCCTCCGCTGTAAGTAAAATACAA
220674

Qy 61 ACTGTTGTCACTTTCCGAGGAGAACAA 87
|||| |||| || | || | ||
Db 220675 ACTGCTGTCCTCCTCTGGAGATGAAAA 220701

RESULT 26

HUMCFTR10

LOCUS HUMCFTR10 203 bp DNA linear PRI 01-NOV-1994

DEFINITION Human cystic fibrosis transmembrane conductance regulator (CFTR)
gene, exon 10.

ACCESSION M55034

VERSION M55034.1 GI:180298

KEYWORDS cystic fibrosis; transmembrane conductance regulator.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 203)

AUTHORS Kerem,B.-S., Zielenski,J., Markiewicz,D., Bozon,D., Gazit,E.,
Yahav,J., Kennedy,D., Riordan,J.R., Collins,F.S., Rommens,J.M. and
Tsui,L.-C.

TITLE Identification of mutations in regions corresponding to the two
putative nucleotide (ATP)-binding folds of the cystic fibrosis gene

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 87 (21), 8447-8451 (1990)

MEDLINE 91046014

PUBMED 2236053

COMMENT Original source text: Human DNA.

FEATURES Location/Qualifiers

source 1..203
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="7q31-q32"

gene 8..199
/gene="CFTR"

exon 8..199
/gene="CFTR"
/note="G00-120-584; putative"
/number=10

variation 130..131
/gene="CFTR"
/note="G00-120-584; putative"

/replace="tatca"
ORIGIN Chromosome 7q31-q32.

Query Match 31.6%; Score 32.2; DB 9; Length 203;
Best Local Similarity 63.6%; Pred. No. 5.7;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || || |||| || | | | || || | || || ||
Db 28 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTCATTCTG 87

Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
Db 88 TTCTCAGTTTTCTGGA 104

RESULT 27

HUMCFTR1

LOCUS HUMCFTR1 206 bp DNA linear PRI 26-SEP-2002

DEFINITION Homo sapiens cystic fibrosis transmembrane conductance regulator
 (CFTR) gene, exon 10.

ACCESSION M55025

VERSION M55025.1 GI:180297

KEYWORDS cystic fibrosis; transmembrane conductance regulator.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 206)

AUTHORS Kerem,B.-S., Zielenski,J., Markiewicz,D., Bozon,D., Gazit,E.,
 Yahav,J., Kennedy,D., Riordan,J.R., Collins,F.S., Rommens,J.M. and
 Tsui,L.-C.

TITLE Identification of mutations in regions corresponding to the two
 putative nucleotide (ATP)-binding folds of the cystic fibrosis gene

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 87 (21), 8447-8451 (1990)

MEDLINE 91046014

PUBMED 2236053

FEATURES Location/Qualifiers

source

1. .206
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q31-q32"

variation

7
/gene="CFTR"
/note="G00-120-584; putative; 1717-"
/replace="a"

gene

8. .199
/gene="CFTR"
/note="cystic fibrosis transmembrane conductance
regulator"

exon

8. .199
/gene="CFTR"
/note="G00-120-584; putative"
/number=10

ORIGIN

Query Match 31.6%; Score 32.2; DB 9; Length 206;
Best Local Similarity 63.6%; Pred. No. 5.7;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
          ||| ||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| |||
Db      28 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTCATTCTG 87

Qy      65 TTGTCACTTTCCGAGGA 81
          || ||| ||| | |||
Db      88 TTCTCAGTTTTCCTGGA 104
```

RESULT 28

MFCFTRW11

LOCUS MFCFTRW11 261 bp DNA linear PRI 01-JUL-2000

DEFINITION Macaca fascicularis cystic fibrosis transmembrane conductance
regulator (CFTR) gene, exon 10.

ACCESSION AF162161

VERSION AF162161.1 GI:8886448

KEYWORDS .

SEGMENT 11 of 27

SOURCE Macaca fascicularis (crab-eating macaque)

ORGANISM Macaca fascicularis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
Cercopithecinae; Macaca.

REFERENCE 1 (bases 1 to 261)

AUTHORS Wine, J.J., Kuo, E., Hurlock, G., Glavac, D. and Dean, M.

TITLE Genomic sequence of CFTR in five primate species

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 261)

AUTHORS Wine, J.J., Kuo, E., Hurlock, G., Glavac, D. and Dean, M.

TITLE Direct Submission

JOURNAL Submitted (24-JUN-1999) Psychology, Stanford University, Building
420, Main Quad, Stanford, CA 94305-2130, USA

FEATURES Location/Qualifiers

```
source      1..261
              /organism="Macaca fascicularis"
              /mol_type="genomic DNA"
              /db_xref="taxon:9541"
exon        31..222
              /gene="CFTR"
              /number=10
```

ORIGIN

Query Match 31.6%; Score 32.2; DB 9; Length 261;
Best Local Similarity 63.6%; Pred. No. 5.7;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
          ||| ||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| |||
Db      51 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTCATTCTG 110

Qy      65 TTGTCACTTTCCGAGGA 81
          || ||| ||| | |||
```

Db 111 TTCTCAGTTTTCTGGA 127

RESULT 29

MFUSCFTR11

LOCUS MFUSCFTR11 261 bp DNA linear PRI 03-AUG-1999

DEFINITION Macaca fuscata cystic fibrosis transmembrane conductance regulator (CFTR) gene, exon 10.

ACCESSION AF162357

VERSION AF162357.1 GI:5679203

KEYWORDS .

SEGMENT 11 of 27

SOURCE Macaca fuscata (Japanese macaque)

ORGANISM Macaca fuscata

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Cercopithecinae; Macaca.

REFERENCE 1 (bases 1 to 261)

AUTHORS Wine,J.J., Kuo,E., Hurlock,G., Glavac,D. and Dean,M.

TITLE CFTR genomic sequences from five primate species

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 261)

AUTHORS Wine,J.J., Kuo,E., Hurlock,G., Glavac,D. and Dean,M.

TITLE Direct Submission

JOURNAL Submitted (24-JUN-1999) Psychology, Stanford University, Building 420, Main Quad, Stanford, CA 94305-2130, USA

FEATURES Location/Qualifiers

source 1..261
/organism="Macaca fuscata"
/mol_type="genomic DNA"
/db_xref="taxon:9542"
exon 31..222
/gene="CFTR"
/number=10

ORIGIN

Query Match 31.6%; Score 32.2; DB 9; Length 261;

Best Local Similarity 63.6%; Pred. No. 5.7;

Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| ||||| || | | | ||| || | ||| | || | |||

Db 51 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTATTCTG 110

Qy 65 TTGTCACCTTCCGAGGA 81
|| ||| ||| | |||

Db 111 TTCTCAGTTTTCTGGA 127

RESULT 30

MNCFTR11

LOCUS MNCFTR11 261 bp DNA linear PRI 03-AUG-1999

DEFINITION Macaca nemestrina cystic fibrosis transmembrane conductance regulator (CFTR) gene, exon 10.

ACCESSION AF162384

VERSION AF162384.1 GI:5679232

KEYWORDS .

```

SEGMENT      11 of 27
SOURCE       Macaca nemestrina (pig-tailed macaque)
ORGANISM      Macaca nemestrina
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
              Cercopithecinae; Macaca.
REFERENCE    1 (bases 1 to 261)
AUTHORS      Wine,J.J., Kuo,E., Hurlock,G., Glavac,D. and Dean,M.
TITLE        CFTR genomic sequences from five primate species
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 261)
AUTHORS      Wine,J.J., Kuo,E., Hurlock,G., Glavac,D. and Dean,M.
TITLE        Direct Submission
JOURNAL      Submitted (24-JUN-1999) Psychology, Stanford University, Building
              420, Main Quad, Stanford, CA 94305-2130, USA
FEATURES
  source      Location/Qualifiers
              1. .261
              /organism="Macaca nemestrina"
              /mol_type="genomic DNA"
              /db_xref="taxon:9545"
  exon        31. .222
              /gene="CFTR"
              /number=10
ORIGIN
Query Match      31.6%;  Score 32.2;  DB 9;  Length 261;
Best Local Similarity  63.6%;  Pred. No. 5.7;
Matches  49;  Conservative  0;  Mismatches  28;  Indels  0;  Gaps  0;

Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
      || | ||| ||      ||| |||| || | | | ||| || | ||| | || |||
Db      51 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG 110

Qy      65 TTGTCACTTTCCGAGGA 81
      || ||| ||| | |||
Db      111 TTCTCAGTTTTCCTGGA 127

RESULT 31
PHACFTR11
LOCUS        PHACFTR11                261 bp    DNA        linear    PRI 03-AUG-1999
DEFINITION   Papio hamadryas anubis cystic fibrosis transmembrane conductance
              regulator (CFTR) gene, exon 10.
ACCESSION    AF162411
VERSION      AF162411.1  GI:5679263
KEYWORDS     .
SEGMENT      11 of 27
SOURCE       Papio anubis (olive baboon)
ORGANISM      Papio anubis
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
              Cercopithecinae; Papio.
REFERENCE    1 (bases 1 to 261)
AUTHORS      Wine,J.J., Kuo,E., Hurlock,G., Glavac,D. and Dean,M.
TITLE        CFTR genomic sequences from five primate species
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 261)

```

AUTHORS Wine,J.J., Kuo,E., Hurlock,G., Glavac,D. and Dean,M.
 TITLE Direct Submission
 JOURNAL Submitted (25-JUN-1999) Psychology, Stanford University, Building
 420, Main Quad, Stanford, CA 94305-2130, USA

FEATURES Location/Qualifiers
 source 1. .261
 /organism="Papio anubis"
 /mol_type="genomic DNA"
 /sub_species="anubis"
 /db_xref="taxon:9555"
 exon 31. .222
 /gene="CFTR"
 /number=10

ORIGIN

Query Match 31.6%; Score 32.2; DB 9; Length 261;
 Best Local Similarity 63.6%; Pred. No. 5.7;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || ||| ||||| || | | | ||| || | ||| | || |||
 Db 51 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG 110

Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 111 TTCTCAGTTTTCCTGGA 127

RESULT 32
 RMCFT11
 LOCUS RMCFT11 261 bp DNA linear PRI 18-APR-1998
 DEFINITION Macaca mulatta cystic fibrosis transmembrane conductance regulator
 (CFTR) gene, exon 10.
 ACCESSION AF016934
 VERSION AF016934.1 GI:3057098
 KEYWORDS .
 SEGMENT 11 of 27
 SOURCE Macaca mulatta (rhesus monkey)
 ORGANISM Macaca mulatta
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
 Cercopithecinae; Macaca.

REFERENCE 1 (bases 1 to 261)
 AUTHORS Wine,J.J., Glavac,D., Hurlock,G., Robinson,C., Lee,M., Potocnik,U.,
 Ravnik-Glavac,M. and Dean,M.
 TITLE Genomic DNA sequence of Rhesus (M. mulatta) cystic fibrosis (CFTR)
 gene
 JOURNAL Mamm. Genome 9 (4), 301-305 (1998)
 MEDLINE 98191731
 PUBMED 9530627

REFERENCE 2 (bases 1 to 261)
 AUTHORS Wine,J.J., Glavac,D., Hurlock,G., Robinson,C., Lee,M., Potocnik,U.,
 Ravnik-Glavac,M. and Dean,M.
 TITLE Direct Submission
 JOURNAL Submitted (04-AUG-1997) Psychology, Stanford University, Bldg. 420
 (Jordan Hall), Stanford, CA 94305-2103, USA

FEATURES Location/Qualifiers

```

source      1. .261
             /organism="Macaca mulatta"
             /mol_type="genomic DNA"
             /db_xref="taxon:9544"
intron      <1. .30
             /gene="CFTR"
             /number=9
exon        31. .222
             /gene="CFTR"
             /number=10
intron      223. .>261
             /gene="CFTR"
             /number=10

```

ORIGIN

```

Query Match      31.6%; Score 32.2; DB 9; Length 261;
Best Local Similarity 63.6%; Pred. No. 5.7;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
        ||| ||||| || ||| ||||| || ||| ||| || ||| ||| |||
Db      51 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGGAGAATTTTCATTCTG 110

Qy      65 TTGTCACTTTCCGAGGA 81
        || ||| ||| | |||
Db      111 TTCTCAGTTTTCCTGGA 127

```

RESULT 33

AR166291

LOCUS AR166291 420 bp DNA linear PAT 17-OCT-2001

DEFINITION Sequence 64 from patent US 6280978.

ACCESSION AR166291

VERSION AR166291.1 GI:16241555

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 420)

AUTHORS Mitchell,L.G. and Garcia-Blanco,M.A.

TITLE Methods and compositions for use in spliceosome mediated RNA trans-splicing

JOURNAL Patent: US 6280978-A 64 28-AUG-2001;

FEATURES Location/Qualifiers

```

source      1. .420
             /organism="unknown"
             /mol_type="unassigned DNA"

```

ORIGIN

```

Query Match      31.6%; Score 32.2; DB 6; Length 420;
Best Local Similarity 63.6%; Pred. No. 5.7;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
        ||| ||||| || ||| ||||| || ||| ||| || ||| ||| |||
Db      128 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGGAGAATTTTCATTCTG 187

```

Qy 65 TTGTCACCTTTCCGAGGA 81
 || ||| ||| | |||
 Db 188 TTCTCAGTTTTCCTGGA 204

RESULT 34

AR381208

LOCUS AR381208 795 bp DNA linear PAT 18-DEC-2003

DEFINITION Sequence 9 from patent US 6607911.

ACCESSION AR381208

VERSION AR381208.1 GI:40088995

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 795)

AUTHORS Gordon,J. and Rundell,C.A.

TITLE Compositions and methods relating to control DNA construct

JOURNAL Patent: US 6607911-A 9 19-AUG-2003;

FEATURES Location/Qualifiers

source 1. .795

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match 31.6%; Score 32.2; DB 6; Length 795;

Best Local Similarity 63.6%; Pred. No. 5.7;

Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || ||| ||||| || | | | ||| || | ||| | |||
 Db 369 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTCATTCTG 428
 Qy 65 TTGTCACCTTTCCGAGGA 81
 || ||| ||| | |||
 Db 429 TTCTCAGTTTTCCTGGA 445

RESULT 35

HUMCFTRA10

LOCUS HUMCFTRA10 831 bp DNA linear PRI 10-JAN-2001

DEFINITION Human cystic fibrosis transmembrane conductance regulator (CFTR) gene, exon 10.

ACCESSION M55115

VERSION M55115.1 GI:306520

KEYWORDS cystic fibrosis transmembrane conductance regulator.

SEGMENT 10 of 26

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 831)

AUTHORS Zielenski,J., Rozmahel,R., Bozon,D., Kerem,B., Grzelczak,Z.,

Riordan,J.R., Rommens,J. and Tsui,L.C.

TITLE Genomic DNA sequence of the cystic fibrosis transmembrane conductance regulator (CFTR) gene

JOURNAL Genomics 10 (1), 214-228 (1991)

MEDLINE 91257831
PUBMED 1710598

FEATURES Location/Qualifiers
source 1. .831
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="7q31-q32"
exon 308. .499
/gene="CFTR"
/note="G00-120-584"
/number=10

ORIGIN

Query Match 31.6%; Score 32.2; DB 9; Length 831;
Best Local Similarity 63.6%; Pred. No. 5.7;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
||| ||||| || ||| ||||| || | | | ||| || | ||| |||
Db 328 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTCAATCTG 387

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 388 TTCTCAGTTTTCCTGGA 404

RESULT 36

G18240

LOCUS G18240 831 bp DNA linear STS 28-SEP-1998

DEFINITION sWSS853 Eric D. Green Homo sapiens STS genomic, sequence tagged site.

ACCESSION G18240

VERSION G18240.1 GI:1222697

KEYWORDS STS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 831)

AUTHORS Bouffard,G.G., Iyer,L.M., Idol,J.R., Braden,V.V., Cunningham,A.F., Weintraub,L.A., Mohr-Tidwell,R.M., Peluso,D.C., Fulton,R.S., Leckie,M.P. and Green,E.D.

TITLE A collection of 1814 human chromosome 7-specific STSs

JOURNAL Genome Res. 7 (1), 59-64 (1997)

MEDLINE 97189344

PUBMED 9037602

REFERENCE 2 (bases 1 to 831)

AUTHORS Green,E.D.

TITLE Human chromosome 7 STSs (1997)

JOURNAL Unpublished (1997)

COMMENT Synonyms: CFTR

GDB: GDB:3754054

GDB_DSEG: CFTR

Contact: Eric D. Green

Genome Technology Branch

National Human Genome Research Institute/NIH

49 Convent Dr., MSC4431, Bldg. 49, Rm. 2A08, Bethesda, MD 20892
Tel: 3014020201
Fax: 3014024735
Email: egreen@nhgri.nih.gov
Primer A: CAGTTTTCCTGGATTATGCCTGG
Primer B: GTTGGCATGCTTTGATGACGCTTC
STS size: 100
PCR Profile:

 Presoak: 0 degrees C for 0.00 minute(s)
 Denaturation: 92 degrees C for 1.00 minute(s)
 Annealing: 62 degrees C for 2.00 minute(s)
 Polymerization: 72 degrees C for 2.00 minute(s)
 PCR Cycles: 35
 Thermal Cyclers: PerkinElmer TC

Protocol:

 Template: 30-100 ng
 Primer: each 1 uM
 dNTPs: each 200 uM
 Taq Polymerase: 0.05 units/ul
 Total Vol: 5 ul

Buffer:

 MgCl2: 2.5 mM
 KCl: 50 mM
 Tris-HCl: 10 mM
 pH: 8.3

 This STS was developed from sequence determined by another investigator. See GenBank record: M55115 For additional information about the NHGRI chromosome 7 mapping project, see <http://www.nhgri.nih.gov/DIR/DTB/CHR7>. Also see Genomics 11:548-64 (1991) [MUID=92128937].

FEATURES	Location/Qualifiers
source	1. .831 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /map="7" /clone_lib="Eric D. Green"
gene	1. .831 /gene="CFTR"
STS	392. .491 /gene="CFTR"
primer_bind	392. .414 /gene="CFTR"
primer_bind	complement(468. .491)

ORIGIN

Query Match 31.6%; Score 32.2; DB 11; Length 831;
Best Local Similarity 63.6%; Pred. No. 5.7;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy	5	TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG	64
Db	328	TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG	387
Qy	65	TTGTCACTTTCCGAGGA	81

Db || ||| ||| | |||
388 TTCTCAGTTTTCTGGA 404

RESULT 37

AR076451

LOCUS AR076451 2640 bp DNA linear PAT 30-AUG-2000

DEFINITION Sequence 1 from patent US 5958893.

ACCESSION AR076451

VERSION AR076451.1 GI:10003197

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 2640)

AUTHORS Welsh,M.J. and Sheppard,D.N.

TITLE Genes and proteins for treating cystic fibrosis

JOURNAL Patent: US 5958893-A 1 28-SEP-1999;

FEATURES Location/Qualifiers

source 1. .2640

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Query Match 31.6%; Score 32.2; DB 6; Length 2640;

Best Local Similarity 63.6%; Pred. No. 5.8;

Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64

 || | |||| || ||| |||| || | | | ||| || | ||| | || |||

Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTATTCTG 1604

Qy 65 TTGTCACTTTCCGAGGA 81

 || ||| ||| | |||

Db 1605 TTCTCAGTTTTCTGGA 1621

RESULT 38

I46970

LOCUS I46970 2640 bp DNA linear PAT 07-OCT-1997

DEFINITION Sequence 1 from patent US 5639661.

ACCESSION I46970

VERSION I46970.1 GI:2470935

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 2640)

AUTHORS Welsh,M.J. and Sheppard,D.N.

TITLE Genes and proteins for treating cystic fibrosis

JOURNAL Patent: US 5639661-A 1 17-JUN-1997;

FEATURES Location/Qualifiers

source 1. .2640

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Query Match 31.6%; Score 32.2; DB 6; Length 2640;
 Best Local Similarity 63.6%; Pred. No. 5.8;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 ||| ||||| || ||| ||||| || | ||| ||| ||| ||| ||| |||
 Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTCATTCTG 1604

Qy 65 TTGTCACCTTCCGAGGA 81
 || ||| ||| | |||
 Db 1605 TTCTCAGTTTTCTGGA 1621

RESULT 39

HUMCFTR10E

LOCUS HUMCFTR10E 2908 bp DNA linear PRI 21-APR-1996

DEFINITION Homo sapiens cystic fibrosis transmembrane conductance regulator (CFTR) gene, exon 10.

ACCESSION L49160

VERSION L49160.1 GI:1160930

KEYWORDS CFTR gene; cystic fibrosis transmembrane conductance regulator.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2908)

AUTHORS Xu,Z. and Gruenert,D.C.

TITLE Human CFTR gene sequences in regions flanking exon 10: a simple repeat sequence polymorphism in intron 9

JOURNAL Biochem. Biophys. Res. Commun. 219 (1), 140-145 (1996)

MEDLINE 96190683

PUBMED 8619797

COMMENT Original source text: Homo sapiens (clone: T6/20) DNA.

FEATURES Location/Qualifiers

source 1. .2908
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /map="7q31-q32"
 /clone="T6/20"

gene 1. .2908
 /gene="CFTR"

intron <1. .1055
 /gene="CFTR"
 /note="G00-120-584; does not fit consensus"
 /number=9

exon 1056. .1256
 /gene="CFTR"
 /note="G00-120-584"
 /number=10

intron 1257. .>2908
 /gene="CFTR"
 /note="G00-120-584"
 /number=10

ORIGIN

Query Match 31.6%; Score 32.2; DB 9; Length 2908;
Best Local Similarity 63.6%; Pred. No. 5.8;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| ||||| || | | | ||| || | ||| |||
Db 1085 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTCATTCTG 1144

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 1145 TTCTCAGTTTTCTGGA 1161

RESULT 40

AR240920
LOCUS AR240920 4443 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 1 from patent US 6468793.
ACCESSION AR240920
VERSION AR240920.1 GI:27286127
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 4443)
AUTHORS Teem, J.L.
TITLE CFTR genes and proteins for cystic fibrosis gene therapy
JOURNAL Patent: US 6468793-A 1 22-OCT-2002;
FEATURES Location/Qualifiers
source 1. .4443
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 31.6%; Score 32.2; DB 6; Length 4443;
Best Local Similarity 63.6%; Pred. No. 5.8;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| ||||| || | | | ||| || | ||| |||
Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTCATTCTG 1472

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 1473 TTCTCAGTTTTCTGGA 1489

RESULT 41

AR240921
LOCUS AR240921 4443 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 3 from patent US 6468793.
ACCESSION AR240921
VERSION AR240921.1 GI:27286128
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 4443)

AUTHORS Teem, J.L.
TITLE CFTR genes and proteins for cystic fibrosis gene therapy
JOURNAL Patent: US 6468793-A 3 22-OCT-2002;
FEATURES Location/Qualifiers
source 1. .4443
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 31.6%; Score 32.2; DB 6; Length 4443;
Best Local Similarity 63.6%; Pred. No. 5.8;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| ||||| || | | ||| || | ||| || |||
Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTTCATTCTG 1472
Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 1473 TTCTCAGTTTTCCTGGA 1489

RESULT 42

AR240922

LOCUS AR240922 4443 bp DNA linear PAT 20-DEC-2002

DEFINITION Sequence 5 from patent US 6468793.

ACCESSION AR240922

VERSION AR240922.1 GI:27286129

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 4443)

AUTHORS Teem, J.L.

TITLE CFTR genes and proteins for cystic fibrosis gene therapy

JOURNAL Patent: US 6468793-A 5 22-OCT-2002;

FEATURES Location/Qualifiers

source 1. .4443
/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match 31.6%; Score 32.2; DB 6; Length 4443;
Best Local Similarity 63.6%; Pred. No. 5.8;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| ||||| || | | ||| || | ||| || |||
Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTTCATTCTG 1472
Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 1473 TTCTCAGTTTTCCTGGA 1489

RESULT 43

AR240923

LOCUS AR240923 4443 bp DNA linear PAT 20-DEC-2002
 DEFINITION Sequence 7 from patent US 6468793.
 ACCESSION AR240923
 VERSION AR240923.1 GI:27286130
 KEYWORDS .
 SOURCE Unknown.
 ORGANISM Unknown.
 Unclassified.
 REFERENCE 1 (bases 1 to 4443)
 AUTHORS Teem, J.L.
 TITLE CFTR genes and proteins for cystic fibrosis gene therapy
 JOURNAL Patent: US 6468793-A 7 22-OCT-2002;
 FEATURES Location/Qualifiers
 source 1. .4443
 /organism="unknown"
 /mol_type="genomic DNA"
 ORIGIN

Query Match 31.6%; Score 32.2; DB 6; Length 4443;
 Best Local Similarity 63.6%; Pred. No. 5.8;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || ||| |||| || | | |||| || | ||| || ||||
 Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGGAGAATTTTCATTCTG 1472
 Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 1473 TTCTCAGTTTTCTGGA 1489

RESULT 44
 AR240924

LOCUS AR240924 4443 bp DNA linear PAT 20-DEC-2002
 DEFINITION Sequence 9 from patent US 6468793.
 ACCESSION AR240924
 VERSION AR240924.1 GI:27286131
 KEYWORDS .
 SOURCE Unknown.
 ORGANISM Unknown.
 Unclassified.
 REFERENCE 1 (bases 1 to 4443)
 AUTHORS Teem, J.L.
 TITLE CFTR genes and proteins for cystic fibrosis gene therapy
 JOURNAL Patent: US 6468793-A 9 22-OCT-2002;
 FEATURES Location/Qualifiers
 source 1. .4443
 /organism="unknown"
 /mol_type="genomic DNA"
 ORIGIN

Query Match 31.6%; Score 32.2; DB 6; Length 4443;
 Best Local Similarity 63.6%; Pred. No. 5.8;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || ||| |||| || | | |||| || | ||| || ||||

Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTCATTCTG 1472

Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||

Db 1473 TTCTCAGTTTTCTGGA 1489

RESULT 45

AR240925

LOCUS AR240925 4443 bp DNA linear PAT 20-DEC-2002

DEFINITION Sequence 11 from patent US 6468793.

ACCESSION AR240925

VERSION AR240925.1 GI:27286132

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 4443)

AUTHORS Teem, J.L.

TITLE CFTR genes and proteins for cystic fibrosis gene therapy

JOURNAL Patent: US 6468793-A 11 22-OCT-2002;

FEATURES Location/Qualifiers

source 1..4443
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Query Match 31.6%; Score 32.2; DB 6; Length 4443;

Best Local Similarity 63.6%; Pred. No. 5.8;

Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || ||| |||| || | | ||| || | ||| | |||

Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTCATTCTG 1472

Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||

Db 1473 TTCTCAGTTTTCTGGA 1489

RESULT 46

AR240926

LOCUS AR240926 4443 bp DNA linear PAT 20-DEC-2002

DEFINITION Sequence 13 from patent US 6468793.

ACCESSION AR240926

VERSION AR240926.1 GI:27286133

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 4443)

AUTHORS Teem, J.L.

TITLE CFTR genes and proteins for cystic fibrosis gene therapy

JOURNAL Patent: US 6468793-A 13 22-OCT-2002;

FEATURES Location/Qualifiers

source 1..4443
 /organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match 31.6%; Score 32.2; DB 6; Length 4443;
Best Local Similarity 63.6%; Pred. No. 5.8;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
||| ||||| || ||| ||||| || | | | ||| || | ||| ||| |||
Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG 1472
Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 1473 TTCTCAGTTTTCCTGGA 1489

RESULT 47

AR240927

LOCUS AR240927 4443 bp DNA linear PAT 20-DEC-2002

DEFINITION Sequence 15 from patent US 6468793.

ACCESSION AR240927

VERSION AR240927.1 GI:27286134

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 4443)

AUTHORS Teem, J.L.

TITLE CFTR genes and proteins for cystic fibrosis gene therapy

JOURNAL Patent: US 6468793-A 15 22-OCT-2002;

FEATURES Location/Qualifiers

source 1. .4443

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match 31.6%; Score 32.2; DB 6; Length 4443;
Best Local Similarity 63.6%; Pred. No. 5.8;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
||| ||||| || ||| ||||| || | | | ||| || | ||| ||| |||
Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG 1472
Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 1473 TTCTCAGTTTTCCTGGA 1489

RESULT 48

AR240928

LOCUS AR240928 4443 bp DNA linear PAT 20-DEC-2002

DEFINITION Sequence 17 from patent US 6468793.

ACCESSION AR240928

VERSION AR240928.1 GI:27286135

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.
 Unclassified.
 REFERENCE 1 (bases 1 to 4443)
 AUTHORS Teem,J.L.
 TITLE CFTR genes and proteins for cystic fibrosis gene therapy
 JOURNAL Patent: US 6468793-A 17 22-OCT-2002;
 FEATURES Location/Qualifiers
 source 1. .4443
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Query Match 31.6%; Score 32.2; DB 6; Length 4443;
 Best Local Similarity 63.6%; Pred. No. 5.8;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 ||| ||||| || ||| ||||| || | | | ||| || | ||| ||| |||
 Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG 1472
 Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 1473 TTCTCAGTTTTCTGGA 1489

RESULT 49

AR240929

LOCUS AR240929 4443 bp DNA linear PAT 20-DEC-2002

DEFINITION Sequence 19 from patent US 6468793.

ACCESSION AR240929

VERSION AR240929.1 GI:27286136

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 4443)

AUTHORS Teem,J.L.

TITLE CFTR genes and proteins for cystic fibrosis gene therapy

JOURNAL Patent: US 6468793-A 19 22-OCT-2002;

FEATURES Location/Qualifiers

source 1. .4443
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Query Match 31.6%; Score 32.2; DB 6; Length 4443;
 Best Local Similarity 63.6%; Pred. No. 5.8;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 ||| ||||| || ||| ||||| || | | | ||| || | ||| ||| |||
 Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG 1472
 Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 1473 TTCTCAGTTTTCTGGA 1489

RESULT 50

AX111569

LOCUS AX111569 4443 bp DNA linear PAT 30-APR-2001

DEFINITION Sequence 3 from Patent WO0125421.

ACCESSION AX111569

VERSION AX111569.1 GI:13927859

KEYWORDS .

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Teem, J.L.

TITLE Materials and method for detecting interaction of cftr polypeptides

JOURNAL Patent: WO 0125421-A 3 12-APR-2001;

Florida State University Research Foundation (US)

FEATURES Location/Qualifiers

source 1. .4443

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 31.6%; Score 32.2; DB 6; Length 4443;

Best Local Similarity 63.6%; Pred. No. 5.8;

Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAGACTG 64

|| | |||| || ||| |||| || | | | || | || | || | ||

Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTCATTCTG 1472

Qy 65 TTGTCACTTTCCGAGGA 81

|| ||| ||| | |||

Db 1473 TTCTCAGTTTCTCTGGA 1489

Search completed: April 29, 2004, 17:05:53

Job time : 440.147 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 29, 2004, 14:53:09 ; Search time 49.1699 Seconds
(without alignments)
8812.639 Million cell updates/sec

Title: US-09-989-981A-9_COPY_3_104
Perfect score: 102
Sequence: 1 ctggtaggtgagatctctga.....aacaagctgtcctggaggcc 102

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		%								
Result		Query								
No.	Score	Match	Length	DB	ID	Description				
c	1	102	100.0	2019	7	AAD48881	Aad48881 Mouse ABC			
c	2	102	100.0	2564	6	ABN90022	Abn90022 Mouse clo			
	3	102	100.0	6043	7	AAD48884	Aad48884 ABCG5-ABC			
c	4	87.6	85.9	2669	7	AAD48883	Aad48883 Human ABC			
	5	32.2	31.6	180	4	ABA71163	Aba71163 Human foe			
	6	32.2	31.6	180	4	AAI51393	Aai51393 Probe #20			
	7	32.2	31.6	180	4	AAK45448	Aak45448 Human bon			

8	32.2	31.6	180	4	AAK19459	Aak19459	Human bra
9	32.2	31.6	180	4	ABS45131	Abs45131	Human liv
10	32.2	31.6	180	6	ABS19713	Abs19713	Human gen
11	32.2	31.6	494	4	ABA58823	Aba58823	Human foe
12	32.2	31.6	494	4	AAI38528	Aai38528	Probe #72
13	32.2	31.6	494	4	AAK32713	Aak32713	Human bon
14	32.2	31.6	494	4	AAK06977	Aak06977	Human bra
15	32.2	31.6	494	4	ABS32432	Abs32432	Human liv
16	32.2	31.6	494	6	ABS07509	Abs07509	Human gen
17	32.2	31.6	500	3	AAZ99413	Aaz99413	Trans-spl
18	32.2	31.6	500	6	ABQ73502	Abq73502	Pre-trans
19	32.2	31.6	795	7	ABZ24468	Abz24468	Cystic fi
20	32.2	31.6	831	9	ADE77694	Ade77694	Human cys
21	32.2	31.6	2640	2	AAT04005	Aat04005	Truncated
22	32.2	31.6	3069	6	ABQ73521	Abq73521	Mouse fac
23	32.2	31.6	4443	4	AAF84742	Aaf84742	DNA encod
24	32.2	31.6	4443	8	ABX16100	Abx16100	Human cDN
25	32.2	31.6	4443	8	ABX16094	Abx16094	Human cDN
26	32.2	31.6	4443	8	ABX16099	Abx16099	Human cDN
27	32.2	31.6	4443	8	ABX16097	Abx16097	Human cDN
28	32.2	31.6	4443	8	ABX16103	Abx16103	Human cDN
29	32.2	31.6	4443	8	ABX16095	Abx16095	Human cDN
30	32.2	31.6	4443	8	ABX16098	Abx16098	Human cDN
31	32.2	31.6	4443	8	ABX16102	Abx16102	Human cDN
32	32.2	31.6	4443	8	ABX16096	Abx16096	Human cDN
33	32.2	31.6	4443	8	ABX16101	Abx16101	Human cDN
34	32.2	31.6	4560	2	AAZ11643	Aaz11643	CFTR prot
35	32.2	31.6	4845	5	AAS81827	Aas81827	DNA encod
36	32.2	31.6	4894	2	AAQ13605	Aaq13605	Cystic fi
37	32.2	31.6	5635	2	AAQ68002	Aaq68002	Ad2/CFTR-
38	32.2	31.6	6126	2	AAQ13053	Aaq13053	CFTR delt
39	32.2	31.6	6126	2	AAX35553	Aax35553	DeltaF508
40	32.2	31.6	6126	6	AAS20529	Aas20529	Human del
41	32.2	31.6	6126	8	ADA37386	Ada37386	DNA encod
42	32.2	31.6	6127	2	AAQ11371	Aaq11371	Mutant cy
43	32.2	31.6	6128	2	AAQ13068	Aaq13068	CFTR 556
44	32.2	31.6	6128	2	AAQ13072	Aaq13072	CFTR 3659
45	32.2	31.6	6129	2	AAQ13056	Aaq13056	CFTR G178
46	32.2	31.6	6129	2	AAQ13060	Aaq13060	CFTR S549
47	32.2	31.6	6129	2	AAQ13071	Aaq13071	CFTR 1717
48	32.2	31.6	6129	2	AAQ13054	Aaq13054	CFTR G85E
49	32.2	31.6	6129	2	AAQ13065	Aaq13065	CFTR L107
50	32.2	31.6	6129	2	AAQ13063	Aaq13063	CFTR Y563

ALIGNMENTS

RESULT 1

AAD48881/c

ID AAD48881 standard; DNA; 2019 BP.

XX

AC AAD48881;

XX

DT 24-MAR-2003 (first entry)

XX

DE Mouse ABCG8 DNA.

XX
KW ABC family cholesterol transporter; ABCG8; sterol-related disorder;
KW sitosterolaemia; hyperlipidaemia; hypercholesterolaemia; gall stone;
KW HDL deficiency; atherosclerosis; nutritional deficiency; gene therapy;
KW mouse; ATP-binding cassette; sitosterolaemia susceptibility gene; SSG;
KW ABCG5; gene; ds.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 1. .2019
FT /*tag= a
FT /product= "mABCG8 protein"
FT /transl_except= (pos:1318. .1320, aa:Leu)
XX
PN WO200281691-A2.
XX
PD 17-OCT-2002.
XX
PF 20-NOV-2001; 2001WO-US043823.
XX
PR 20-NOV-2000; 2000US-0252235P.
PR 28-NOV-2000; 2000US-0253645P.
XX
PA (TULA-) TULARIK INC.
PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Hobbs HH, Shan B, Barnes R, Tian H;
XX
DR WPI; 2003-058548/05.
DR P-PSDB; AAE31703.
XX
PT New ABCG8 polypeptides and nucleic acids, useful for treating sterol-
PT related disorders e.g. sitosterolemia, hypercholesterolemia,
PT hyperlipidemia, gall stones, HDL deficiency, atherosclerosis, or
PT nutritional deficiencies.
XX
PS Claim 13; Page 75; 94pp; English.
XX
CC The invention relates to ATP-binding cassette (ABC) family cholesterol
CC transporter, ABCG8 polypeptides and polynucleotides. The invention also
CC provides ABCG5 polypeptides and polynucleotides. ABCG5 gene is also known
CC as sitosterolaemia susceptibility gene (SSG). Sequences of the invention
CC are useful for treating or preventing sterol-related disorders such as
CC sitosterolaemia, hyperlipidaemia, hypercholesterolaemia, gall stones, HDL
CC deficiency, atherosclerosis and nutritional deficiencies. They are also
CC useful in gene therapy. The present sequence is mouse ABCG8 DNA
XX
SQ Sequence 2019 BP; 444 A; 598 C; 510 G; 467 T; 0 U; 0 Other;

Query Match 100.0%; Score 102; DB 7; Length 2019;
Best Local Similarity 100.0%; Pred. No. 2.7e-25;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
|||||
Db 165 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 106

Qy 61 ACTGTTGTCACTTTCCGAGGAGACAAGCTGTCCTGGAGGCC 102
 |||||
 Db 105 ACTGTTGTCACTTTCCGAGGAGACAAGCTGTCCTGGAGGCC 64

RESULT 2

ABN90022/c

ID ABN90022 standard; cDNA; 2564 BP.

XX

AC ABN90022;

XX

DT 16-AUG-2002 (first entry)

XX

DE Mouse clone IMX3_67 extended sequence.

XX

KW Mouse; antiinflammatory; gene therapy; ileitis; DST; ss; TOGA;
 KW digital sequence tag; total gene expression analysis.

XX

OS Mus musculus.

XX

PN WO200231114-A2.

XX

PD 18-APR-2002.

XX

PF 11-OCT-2001; 2001WO-US032091.

XX

PR 11-OCT-2000; 2000US-0239483P.

XX

PA (DIGI-) DIGITAL GENE TECHNOLOGIES INC.

XX

PI Viney JL, Sims JE, Dubose RF, Baum PR, Hasel KW, Hilbush BS;

XX

DR WPI; 2002-426279/45.

XX

PT New isolated nucleic acid molecules that are associated with ileitis, for
 PT preventing, treating, modulating and diagnosing ileitis in a mammalian
 PT subject.

XX

PS Claim 1; Page 266-268; 273pp; English.

XX

CC The invention relates to a novel isolated nucleic acid molecule
 CC comprising a polynucleotide having one of 90 polynucleotide sequences,
 CC given in the specification. The polynucleotides of the invention have
 CC antiinflammatory activity, and may have a use in gene therapy. The
 CC polynucleotide or a polypeptide encoded by it is used for preventing,
 CC treating, modulating or ameliorating a medical condition such as ileitis.
 CC The polypeptide or polynucleotide is also useful for manufacturing a
 CC medicament for treating ileitis. The sequence represents a an extended
 CC cDNA digital sequence tag obtained from a mouse clone by the TOGA (total
 CC gene expression analysis) method

XX

SQ Sequence 2564 BP; 623 A; 722 C; 638 G; 581 T; 0 U; 0 Other;

Query Match 100.0%; Score 102; DB 6; Length 2564;

Best Local Similarity 100.0%; Pred. No. 2.9e-25;

Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      202 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 143

Qy      61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 102
        ||||||||||||||||||||||||||||||||||||
Db      142 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 101

```

RESULT 3

AAD48884

ID AAD48884 standard; DNA; 6043 BP.

XX

AC AAD48884;

XX

DT 24-MAR-2003 (first entry)

XX

DE ABCG5-ABCG8 DNA.

XX

KW ABC family cholesterol transporter; ABCG8; sterol-related disorder;
 KW sitosterolaemia; hyperlipidaemia; hypercholesterolaemia; gall stone;
 KW HDL deficiency; atherosclerosis; nutritional deficiency; gene therapy;
 KW ATP-binding cassette; sitosterolaemia susceptibility gene; SSG; ABCG5;
 KW ds.

XX

OS Unidentified.

XX

FH	Key	Location/Qualifiers
FT	exon	complement(3. .104)
FT		/*tag= a
FT		/number= 2
FT		/note= "Corresponds to ABCG8 gene"
FT	intron	complement(105. .3435)
FT		/*tag= b
FT		/number= 1
FT		/cons_splice= (5'site:NO, 3'site:NO)
FT		/note= "Corresponds to ABCG8 gene"
FT	misc_feature	complement(1098. .1377)
FT		/*tag= c
FT		/note= "ABCG8 intron1 conserved region"
FT	misc_feature	complement(3250. .3294)
FT		/*tag= d
FT		/note= "ABCG8 intron1 conserved region"
FT	exon	3436. .3498
FT		/*tag= e
FT		/number= 1
FT		/note= "Corresponds to ABCG8 gene"
FT	exon	3858. .4003
FT		/*tag= f
FT		/number= 1
FT		/note= "Corresponds to ABCG5 gene"
FT	intron	4004. .4598
FT		/*tag= g
FT		/number= 1
FT		/note= "Corresponds to ABCG5 gene"
FT	exon	4599. .4720

```

FT          /*tag= h
FT          /number= 2
FT          /note= "Corresponds to ABCG5 gene"
FT  intron  4721. .6043
FT          /*tag= i
FT          /number= 2
FT          /partial
FT          /note= "Corresponds to ABCG5 gene"
XX
PN  WO200281691-A2.
XX
PD  17-OCT-2002.
XX
PF  20-NOV-2001; 2001WO-US043823.
XX
PR  20-NOV-2000; 2000US-0252235P.
PR  28-NOV-2000; 2000US-0253645P.
XX
PA  (TULA-) TULARIK INC.
PA  (TEXA ) UNIV TEXAS SYSTEM.
XX
PI  Hobbs HH,  Shan B,  Barnes R,  Tian H;
XX
DR  WPI; 2003-058548/05.
XX
PT  New ABCG8 polypeptides and nucleic acids, useful for treating sterol-
PT  related disorders e.g. sitosterolemia, hypercholesterolemia,
PT  hyperlipidemia, gall stones, HDL deficiency, atherosclerosis, or
PT  nutritional deficiencies.
XX
PS  Disclosure; Fig 3; 94pp; English.
XX
CC  The invention relates to ATP-binding cassette (ABC) family cholesterol
CC  transporter, ABCG8 polypeptides and polynucleotides. The invention also
CC  provides ABCG5 polypeptides and polynucleotides. ABCG5 gene is also known
CC  as sitosterolaemia susceptibility gene (SSG). Sequences of the invention
CC  are useful for treating or preventing sterol-related disorders such as
CC  sitosterolaemia, hyperlipidaemia, hypercholesterolaemia, gall stones, HDL
CC  deficiency, atherosclerosis and nutritional deficiencies. They are also
CC  useful in gene therapy. The present sequence is ABCG8- ABCG5 DNA
XX
SQ  Sequence 6043 BP; 1378 A; 1509 C; 1497 G; 1654 T; 0 U; 5 Other;

```

```

Query Match          100.0%;  Score 102;  DB 7;  Length 6043;
Best Local Similarity 100.0%;  Pred. No. 3.8e-25;
Matches 102;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

Qy          1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAG 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAG 62

Qy          61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 102
             ||||||||||||||||||||||||||||||||||||
Db          63 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 104

```

RESULT 4

AAD48883/c

ID AAD48883 standard; DNA; 2669 BP.

XX

AC AAD48883;

XX

DT 24-MAR-2003 (first entry)

XX

DE Human ABCG8 DNA.

XX

KW ABC family cholesterol transporter; ABCG8; sterol-related disorder;
KW sitosterolaemia; hyperlipidaemia; hypercholesterolaemia; gall stone;
KW HDL deficiency; atherosclerosis; nutritional deficiency; gene therapy;
KW human; ATP-binding cassette; sitosterolaemia susceptibility gene; SSG;
KW ABCG5; gene; ds.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 100..2121

FT /*tag= a

FT /product= "hABCG8 protein"

XX

PN WO200281691-A2.

XX

PD 17-OCT-2002.

XX

PF 20-NOV-2001; 2001WO-US043823.

XX

PR 20-NOV-2000; 2000US-0252235P.

PR 28-NOV-2000; 2000US-0253645P.

XX

PA (TULA-) TULARIK INC.

PA (TEXA) UNIV TEXAS SYSTEM.

XX

PI Hobbs HH, Shan B, Barnes R, Tian H;

XX

DR WPI; 2003-058548/05.

DR P-PSDB; AAE31705.

XX

PT New ABCG8 polypeptides and nucleic acids, useful for treating sterol-
PT related disorders e.g. sitosterolemia, hypercholesterolemia,
PT hyperlipidemia, gall stones, HDL deficiency, atherosclerosis, or
PT nutritional deficiencies.

XX

PS Claim 13; Page 80; 94pp; English.

XX

CC The invention relates to ATP-binding cassette (ABC) family cholesterol
CC transporter, ABCG8 polypeptides and polynucleotides. The invention also
CC provides ABCG5 polypeptides and polynucleotides. ABCG5 gene is also known
CC as sitosterolaemia susceptibility gene (SSG). Sequences of the invention
CC are useful for treating or preventing sterol-related disorders such as
CC sitosterolaemia, hyperlipidaemia, hypercholesterolaemia, gall stones, HDL
CC deficiency, atherosclerosis and nutritional deficiencies. They are also
CC useful in gene therapy. The present sequence is human ABCG8 DNA

XX

SQ Sequence 2669 BP; 595 A; 768 C; 722 G; 584 T; 0 U; 0 Other;

Query Match 85.9%; Score 87.6; DB 7; Length 2669;
Best Local Similarity 91.2%; Pred. No. 3.1e-20;
Matches 93; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
QY      1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGGTGGACTGACCACTGTAGGTGAAGTACAG 60
      ||||| ||| ||||| ||||| ||||| ||| ||||| ||||| |||||
Db      264 CTGGTAGTTGAGGTCTCTGACCTCCAGGGTGGTGGGCTGGCCACTGTAGGTGAAGTACAG 205

QY      61 ACTGTTGTCACTTTCCGAGGAGAAACAAGCTGTCCTGGAGGCC 102
      ||||| ||||| ||||| ||| ||||| |||
Db      204 GCTGTTGTCACTTTTCAGAGGAGAAACAATCTATCCTGGAGGCC 163
```

RESULT 5

ABA71163

ID ABA71163 standard; DNA; 180 BP.

XX

AC ABA71163;

XX

DT 01-FEB-2002 (first entry)

XX

DE Human foetal liver single exon nucleic acid probe #19468.

XX

KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX

OS Homo sapiens.

XX

PN WO200157277-A2.

XX

PD 09-AUG-2001.

XX

PF 30-JAN-2001; 2001WO-US000669.

XX

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

DR WPI; 2001-483447/52.

XX

PT Human genome-derived single exon nucleic acid probes useful for analyzing

PT gene expression in human fetal liver.

XX

PS Claim 4; SEQ ID NO 19468; 639pp + Sequence Listing; English.

XX

CC The invention relates to a single exon nucleic acid probe for measuring

CC human gene expression in a sample derived from human foetal liver. The

CC single exon nucleic acid probes may be used for predicting, measuring and

CC displaying gene expression in samples derived from human fetal liver. The

CC present sequence is a single exon nucleic acid probe of the invention.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 180 BP; 58 A; 29 C; 41 G; 52 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 4; Length 180;
Best Local Similarity 63.6%; Pred. No. 0.28;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| |||| || | | ||| || | ||| | || |||
Db 9 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTTCATTCTG 68

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 69 TTCTCAGTTTTCCTGGA 85

RESULT 6

AAI51393

ID AAI51393 standard; DNA; 180 BP.

XX

AC AAI51393;

XX

DT 17-OCT-2001 (first entry)

XX

DE Probe #20079 used to measure gene expression in human placenta sample.

XX

KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.

XX

OS Homo sapiens.

XX

PN WO200157272-A2.

XX

PD 09-AUG-2001.

XX

PF 30-JAN-2001; 2001WO-US000663.

XX

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

DR WPI; 2001-488897/53.

XX

PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.

XX

PS Claim 25; SEQ ID NO 20079; 654pp; English.

XX

CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders

XX

SQ Sequence 180 BP; 58 A; 29 C; 41 G; 52 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 4; Length 180;

Best Local Similarity 63.6%; Pred. No. 0.28;

Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAGACTG 64

||| ||||| || ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 9 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTTCATTCTG 68

Qy 65 TTGTCACCTTTCCGAGGA 81

|| ||| ||| ||| ||| |||

Db 69 TTCTCAGTTTTCCTGGA 85

RESULT 7

AAK45448

ID AAK45448 standard; DNA; 180 BP.

XX

AC AAK45448;

XX

DT 06-NOV-2001 (first entry)

XX

DE Human bone marrow expressed single exon probe SEQ ID NO: 20005.

XX

KW Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.

XX

OS Homo sapiens.

XX

PN WO200157276-A2.

XX

PD 09-AUG-2001.

XX

PF 30-JAN-2001; 2001WO-US000668.

XX

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

DR WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
PT
XX
PS Example 4; SEQ ID NO 20005; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention
XX
SQ Sequence 180 BP; 58 A; 29 C; 41 G; 52 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 4; Length 180;
Best Local Similarity 63.6%; Pred. No. 0.28;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 ||| ||||| || ||||| || | ||| ||| ||| ||| ||| |||
 Db 9 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGGAAGAATTTCACTCTG 68

Qy 65 TTGTCACCTTCCGAGGA 81
 || ||| ||| | |||
 Db 69 TTCTCAGTTTTCTGGA 85

RESULT 8

AAK19459

ID AAK19459 standard; DNA; 180 BP.

XX

AC AAK19459;

XX

DT 05-NOV-2001 (first entry)

XX

DE Human brain expressed single exon probe SEQ ID NO: 19450.

XX

KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW ss.

XX

OS Homo sapiens.

XX

PN WO200157275-A2.

XX

PD 09-AUG-2001.

XX

PF 30-JAN-2001; 2001WO-US000667.

XX

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
PS Example 4; SEQ ID NO 19450; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention
XX
SQ Sequence 180 BP; 58 A; 29 C; 41 G; 52 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 4; Length 180;
Best Local Similarity 63.6%; Pred. No. 0.28;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | ||| || ||| ||||| || | | ||| || | ||| | || |||
Db 9 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG 68
Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 69 TTCTCAGTTTTCCTGGA 85

RESULT 9
ABS45131
ID ABS45131 standard; DNA; 180 BP.
XX
AC ABS45131;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver single exon probe, SEQ ID No 20121.
XX
KW Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000664.
XX

PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

DR WPI; 2001-488898/53.

XX

PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.

XX

PS Claim 4; SEQ ID NO 20121; 658pp; English.

XX

CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC sequence information for this patent does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 180 BP; 58 A; 29 C; 41 G; 52 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 4; Length 180;
Best Local Similarity 63.6%; Pred. No. 0.28;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| ||||| || | | | ||| || | ||| | || |||

Db 9 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTTCATTCTG 68

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||

Db 69 TTCTCAGTTTTCCTGGA 85

RESULT 10

ABS19713

ID ABS19713 standard; DNA; 180 BP.

XX

AC ABS19713;

XX

DT 19-AUG-2002 (first entry)

XX

DE Human genome-derived single exon probe ORF from lung SEQ ID No 19704.
 XX
 KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease; open reading frame; ORF.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000665.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples.
 XX
 PS Claim 4; SEQ ID NO 19704; 634pp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of probes
 CC ; the novel set of probes which hybridise at high stringency to a nucleic
 CC acid expressed in the human lung; measuring gene expression in a sample
 CC derived from human lung, comprising (a) contacting the array with a
 CC collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of the
 CC array; identifying exons in a eukaryotic genome, comprising (a)
 CC algorithmically predicting at least one exon from genomic sequences of
 CC the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method

CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene expression
 CC analysis, and for identifying exons in a gene, particularly using human
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 CC Karageners syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 CC present sequence is a single exon probe open reading frame of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 180 BP; 58 A; 29 C; 41 G; 52 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 6; Length 180;
 Best Local Similarity 63.6%; Pred. No. 0.28;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || ||| ||||| || | | | ||| || | ||| | || |||
 Db 9 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTATTCTG 68
 Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 69 TTCTCAGTTTTCTGGA 85

RESULT 11

ABA58823

ID ABA58823 standard; DNA; 494 BP.

XX

AC ABA58823;

XX

DT 01-FEB-2002 (first entry)

XX

DE Human foetal liver single exon nucleic acid probe #7128.

XX

KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX

OS Homo sapiens.

XX

PN WO200157277-A2.

XX

PD 09-AUG-2001.

XX

PF 30-JAN-2001; 2001WO-US000669.

XX

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

DR WPI; 2001-483447/52.

XX

PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.

XX

PS Claim 1; SEQ ID NO 7128; 639pp + Sequence Listing; English.

XX

CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a single exon nucleic acid probe of the invention.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 494 BP; 155 A; 81 C; 92 G; 166 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 4; Length 494;
Best Local Similarity 63.6%; Pred. No. 0.38;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
||| ||||| || ||| ||||| || | ||| ||| ||| ||| ||| |||
Db 280 TATGGGAGAACTGGAGCCTTCAGAGGCTAAATTAAGCACAGTGGAAGAATTTTCATTCTG 339

Qy 65 TTGTCACCTTTCCGAGGA 81
|| ||| ||| | |||
Db 340 TTCTCAGTTTTCCTGGA 356

RESULT 12

AAI38528

ID AAI38528 standard; DNA; 494 BP.

XX

AC AAI38528;

XX

DT 17-OCT-2001 (first entry)

XX

DE Probe #7214 used to measure gene expression in human placenta sample.

XX

KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.

XX

OS Homo sapiens.

XX

PN WO200157272-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000663.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488897/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human placenta.
 XX
 PS Claim 25; SEQ ID NO 7214; 654pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes (SENP).
 CC The present sequence is one such probe. The probes are useful for
 CC producing a microarray for predicting, measuring and displaying gene
 CC expression in samples derived from human placenta. The probes are useful
 CC for antenatal diagnosis of human genetic disorders
 XX
 SQ Sequence 494 BP; 155 A; 81 C; 92 G; 166 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 4; Length 494;
 Best Local Similarity 63.6%; Pred. No. 0.38;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 ||| ||||| || ||| ||||| || | | | ||| || | ||| | ||| |||
 Db 280 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGGAAGAATTTTCATTCTG 339
 Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 340 TTCTCAGTTTTCCTGGA 356

RESULT 13

AAK32713

ID AAK32713 standard; DNA; 494 BP.

XX

AC AAK32713;

XX

DT 06-NOV-2001 (first entry)

XX

DE Human bone marrow expressed single exon probe SEQ ID NO: 7270.

XX

KW Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157276-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000668.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488900/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human bone marrow.
 XX
 PS Example 4; SEQ ID NO 7270; 658pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
 CC the probes of the invention
 XX
 SQ Sequence 494 BP; 155 A; 81 C; 92 G; 166 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 4; Length 494;
 Best Local Similarity 63.6%; Pred. No. 0.38;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || ||| ||||| || | | | ||| || | ||| | || |||
 Db 280 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTCACTCTG 339
 Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 340 TTCTCAGTTTTCCTGGA 356

RESULT 14

AAK06977

ID AAK06977 standard; DNA; 494 BP.

XX

AC AAK06977;

XX

DT 05-NOV-2001 (first entry)
 XX
 DE Human brain expressed single exon probe SEQ ID NO: 6968.
 XX
 KW Human; brain expressed exon; gene expression analysis; probe; microarray;
 KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
 KW ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157275-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000667.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483446/52.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains.
 XX
 PS Example 4; SEQ ID NO 6968; 650pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention
 XX
 SQ Sequence 494 BP; 155 A; 81 C; 92 G; 166 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 4; Length 494;
 Best Local Similarity 63.6%; Pred. No. 0.38;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 ||| ||||| || ||| ||||| || | ||| ||| ||| ||| ||| |||
 Db 280 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGGAAGAATTTTCATTCTG 339
 Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 340 TTCTCAGTTTTCCTGGA 356

RESULT 15

ABS32432

ID ABS32432 standard; DNA; 494 BP.

XX

AC ABS32432;

XX

DT 25-FEB-2003 (first entry)

XX

DE Human liver single exon probe, SEQ ID No 7422.

XX

KW Human; single exon nucleic acid probe; liver; cirrhosis;

KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;

KW coronary heart disease; ss.

XX

OS Homo sapiens.

XX

PN WO200157273-A2.

XX

PD 09-AUG-2001.

XX

PF 30-JAN-2001; 2001WO-US000664.

XX

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

DR WPI; 2001-488898/53.

XX

PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.

XX

PS Claim 1; SEQ ID NO 7422; 658pp; English.

XX

CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC sequence information for this patent does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 494 BP; 155 A; 81 C; 92 G; 166 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 4; Length 494;
Best Local Similarity 63.6%; Pred. No. 0.38;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
        || | ||| ||  || |||| ||  | | | ||| || | ||| | || |||
Db      280 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG 339

Qy      65 TTGTCACTTTCCGAGGA 81
        || ||| ||| | |||
Db      340 TTCTCAGTTTTCCTGGA 356
```

RESULT 16

ABS07509

ID ABS07509 standard; DNA; 494 BP.

XX

AC ABS07509;

XX

DT 19-AUG-2002 (first entry)

XX

DE Human genome-derived single exon probe from lung SEQ ID No 7500.

XX

KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;

KW chronic obstructive pulmonary disease; interstitial lung disease;

KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

KW primary ciliary dyskinesia; pulmonary hypertension;

KW hyaline membrane disease.

XX

OS Homo sapiens.

XX

PN WO200186003-A2.

XX

PD 15-NOV-2001.

XX

PF 30-JAN-2001; 2001WO-US000665.

XX

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

DR WPI; 2002-114183/15.

XX

PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.

XX

PS Claim 1; SEQ ID NO 7500; 634pp; English.

XX

CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC ; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karagenen syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 494 BP; 155 A; 81 C; 92 G; 166 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 6; Length 494;
Best Local Similarity 63.6%; Pred. No. 0.38;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
||| ||||| || ||||| || ||||| ||||| ||||| ||||| ||||| |||||
Db 280 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG 339

Qy 65 TTGTCACTTTCCGAGGA 81
||| ||| ||| ||| |||
Db 340 TTCTCAGTTTTCCTGGA 356

RESULT 17

AAZ99413

ID AAZ99413 standard; DNA; 500 BP.

XX

AC AAZ99413;

XX

DT 03-JUL-2000 (first entry)

XX

DE Trans-spliced product of the CFTR target pre-mRNA and a PTM.

XX

KW Pre-mRNA molecule; gene repair; pre-trans-splicing molecule;

KW gene regulation; targeted cell death;

KW cystic fibrosis trans-membrane regulator gene; ss.

XX

OS Homo sapiens.

XX

PN WO200009734-A2.

XX

PD 24-FEB-2000.

XX

PF 12-AUG-1999; 99WO-US018371.

XX

PR 13-AUG-1998; 98US-00133717.

PR 23-SEP-1998; 98US-00158863.

XX

PA (INTR-) INTRONN HOLDINGS LLC.

XX

PI Mitchell LG, Garcia-Blanco MA;

XX

DR WPI; 2000-224360/19.

XX

PT Novel pre-trans-splicing molecules for use in gene regulation, gene

PT repair and targeted cell death particularly gene repair of cystic

PT fibrosis trans-membrane regulator gene.

XX

PS Example 8; Fig 15; 79pp; English.

XX

CC The specification describes a pre-trans-splicing molecule (PTM) which
 CC contains one or more target binding domains, a 3' splice region
 CC comprising a branch point, a pyrimidine tract and a 3' splice acceptor
 CC site, a spacer region separating the mRNA splice region from the target
 CC binding domain, and a nucleotide sequence to be trans-spliced. The method
 CC is used for the in vivo production of a trans-spliced molecule in a
 CC subset of cells. The PTM is used for producing chimeric mRNA molecule by
 CC contacting it with target pre mRNA which is useful for gene regulation,
 CC gene repair and targeted cell death particularly repair of cystic
 CC fibrosis trans-membrane regulator (CFTR) gene. The present sequence
 CC represents the trans-spliced product of the CFTR target pre-mRNA and a
 CC PTM of the invention

XX

SQ Sequence 500 BP; 125 A; 127 C; 102 G; 146 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 3; Length 500;

Best Local Similarity 63.6%; Pred. No. 0.38;

Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || ||| ||||| || | | | ||| || | ||| | || |||
 Db 128 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTTTCATTCTG 187

Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 188 TTCTCAGTTTTCTGGA 204

RESULT 18

ABQ73502

ID ABQ73502 standard; DNA; 500 BP.

XX

AC ABQ73502;

XX

DT 02-OCT-2002 (first entry)

XX

DE Pre-trans-splicing molecule related oligonucleotide #9.

XX

KW Pre-trans-splicing molecule; PTM; spliceosome; cytostatic; gene therapy;

KW immunosuppressive; antimicrobial; gene regulation; gene repair; cancer;

KW targeted cell death; genetic disorder; infectious disorder;

KW autoimmune disease; proliferative disorder; PCR primer; ss.

XX

OS Synthetic.

XX

PN WO200253581-A2.

XX

PD 11-JUL-2002.

XX

PF 08-JAN-2002; 2002WO-US000416.

XX

PR 08-JAN-2001; 2001US-00756095.

PR 08-JAN-2001; 2001US-00756096.

PR 08-JAN-2001; 2001US-00756097.

PR 20-APR-2001; 2001US-00838858.

PR 29-AUG-2001; 2001US-00941492.

XX

PA (INTR-) INTRONN INC.

XX

PI Mitchell LG, Garcia-Blanco MA, Baker CC, Puttaraju M;

PI Mansfield GS, Chao H;

XX

DR WPI; 2002-566693/60.

XX

PT Novel cell having pre-trans-splicing molecules with target binding

PT domains that target binding of PTM to pre-mRNA, 3' or 5' splice region,

PT spacer region, nucleotide sequence to be trans-spliced to target-pre-

PT mRNA.

XX

PS Example; Fig 15A-B; 229pp; English.

XX

CC The present invention describes a cell (I) comprising pre-trans-splicing

CC molecules (PTMs) (II) which have one or more target binding domains (IIa)

CC that target binding of PTM to pre-mRNA, 3' splice region (IIb) that

CC includes branch point pyrimidine tract and 3'splice acceptor site, or 5'

splice site (IIc), spacer region (IIId) that separates RNA splice site from target binding domain, and nucleotide sequence to (IIe) be trans-spliced to target-pre-mRNA. Optionally, the cell comprises (II) either comprising: (A) (IIb) and (IIe); or (B) (IIc), (IIId) and (IIe). The cell may comprise a recombinant vector expressing (II). (I) has cytostatic, immunosuppressive and antimicrobial activities, and can be used in gene therapy. (II) comprising one or more (preferably two or more) (IIa) and (IIb) (or (IIc)), (IIId) and (IIe), or (II) comprising either (A) or (B) (excluding (IIId)), is useful for producing a chimeric RNA molecule in a cell which involves contacting a target pre-mRNA expressed in the cell with (II) that is recognised by nuclear splicing components. The chimeric RNA produced comprises sequences encoding a toxin or translatable protein. The nucleotide sequence to be trans-spliced to target pre-mRNA preferably comprises nucleotide sequences comprising exons 1-10 of cystic fibrosis trans-membrane conductance regulator (CFTR). The chimeric RNA molecule produced using (II) which either comprises (A) or (B) further comprises a nucleotide sequence tag. (I) can be used for gene regulation, gene repair and targeted cell death. (I) can be used for the treatment of various diseases including genetic, infectious or autoimmune diseases and proliferative disorders such as cancer and to regulate gene expression in plants. ABQ73414 to ABQ73536 represent sequences used in the exemplification of the present invention

XX

SQ Sequence 500 BP; 125 A; 128 C; 101 G; 146 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 6; Length 500;

Best Local Similarity 63.6%; Pred. No. 0.38;

Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
      || | ||| || ||| ||| || | | | || | || | || |
Db      128 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG 187

Qy      65 TTGTCACTTTCCGAGGA 81
      || ||| ||| | |||
Db      188 TTCTCAGTTTTCCTGGA 204
```

RESULT 19

ABZ24468

ID ABZ24468 standard; DNA; 795 BP.

XX

AC ABZ24468;

XX

DT 21-MAR-2003 (first entry)

XX

DE Cystic fibrosis transmembrane conductance regulator gene exon 10.

XX

KW Cystic fibrosis transmembrane conductance regulator; CFTR; human;

KW cystic fibrosis; nucleic acid detection; quality assurance; validation;

KW diagnosis; ds.

XX

OS Homo sapiens.

XX

PN WO200296925-A1.

XX

PD 05-DEC-2002.

XX
 PF 24-MAY-2002; 2002WO-US016504.
 XX
 PR 25-MAY-2001; 2001US-00866293.
 XX
 PA (MAIN-) MAINE MEDICAL CENT RES INST.
 PA (MAIN-) MAINE MOLECULAR QUALITY CONTROLS INC.
 XX
 PI Gordon J, Rundell CA;
 XX
 DR WPI; 2003-140437/13.
 XX
 PT Control DNA constructs useful in nucleic acid assays, has vector portion
 PT for expression in a cell and a target nucleic acid comprising fragments
 PT which specify component associated with disease state or environmental
 PT condition.
 XX
 PS Disclosure; Page 74-75; 76pp; English.
 XX
 CC The present sequence is the nucleotide sequence of exon 10 of the human
 CC cystic fibrosis transmembrane conductance regulator (CFTR) gene. Many of
 CC the most common disease-causing mutations are in exon 10 and exon 11 (see
 CC ABZ24469) of the CFTR gene, and genetic screening for these mutations is
 CC therefore advantageous for early diagnosis of cystic fibrosis. The
 CC invention provides control DNA constructs useful in nucleic acid assays.
 CC The DNA constructs have a vector portion for expression in a cell and a
 CC target nucleic acid comprising 2 or more nucleic acid fragments, where
 CC each fragment specifies a component associated with a disease state, an
 CC environmental condition or a biological organism. Each fragment may
 CC comprise at least 1 exon of a gene, and is especially a CFTR exon,
 CC particularly exon 10 and exon 11. The DNA constructs provide controls
 CC useful for quality assurance in the diagnostic detection of complex
 CC genetic diseases such as cystic fibrosis, and for quality assurance in
 CC nucleic acid assays to detect components associated with an environmental
 CC condition or a biological organism
 XX
 SQ Sequence 795 BP; 251 A; 143 C; 135 G; 266 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 7; Length 795;
 Best Local Similarity 63.6%; Pred. No. 0.44;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || ||| ||||| || | | ||| || | ||| | ||
 Db 369 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTCAATTCTG 428
 Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 429 TTCTCAGTTTTCTGGA 445

RESULT 20

ADE77694

ID ADE77694 standard; DNA; 831 BP.

XX

AC ADE77694;

XX

DT 29-JAN-2004 (first entry)
XX
DE Human cystic fibrosis conductance transmembrane regulator exon 10 DNA.
XX
KW ds; human; CFTR; human leukocyte antigen; HLA; genetic testing;
KW carrier screening; genotyping; profiling; polymorphic;
KW multiplexed elongation assay; enzymatic recognition;
KW cystic fibrosis conductance transmembrane regulator;
KW single nucleotide polymorphism; SNP.
XX
OS Homo sapiens.
XX
PN WO2003034029-A2.
XX
PD 24-APR-2003.
XX
PF 15-OCT-2002; 2002WO-US033012.
XX
PR 15-OCT-2001; 2001US-0329427P.
PR 15-OCT-2001; 2001US-0329428P.
PR 15-OCT-2001; 2001US-0329619P.
PR 15-OCT-2001; 2001US-0329620P.
PR 14-MAR-2002; 2002US-0364416P.
XX
PA (BIOA-) BIOARRAY SOLUTIONS LTD.
XX
PI Li AX, Hashmi G, Seul M;
XX
DR WPI; 2003-393553/37.
XX
PT Concurrent interrogation of a number of polymorphic sites, useful for
PT genetic testing, carrier screening, genetic profiling, and identity
PT testing, comprises conducting a multiplexed elongation assay using
PT probes.
XX
PS Example 12; Page 54; 143pp; English.
XX
CC This invention relates to a novel method for the concurrent interrogation
CC of a number of polymorphic sites in the presence of, and without
CC interference from, non-designated polymorphic sites. Specifically, it
CC comprises conducting a multiplexed elongation assay by applying one or
CC more temperature cycles to achieve linear amplification of the target or
CC a combination of annealing and elongation steps under temperature-
CC controlled conditions. Furthermore, this detection method uses probe
CC extension or elongation and relies on enzymatic recognition, a superior
CC technique that no longer depends on differential hybridisation. The
CC present invention describes probes and methods useful for identifying or
CC detecting polymorphisms at one or more designated sites, such that they
CC can identify mutations within the cystic fibrosis conductance
CC transmembrane regulator (CFTR) or the human leukocyte antigen (HLA)
CC genes. In addition, concurrent interrogation of a multiplicity of
CC polymorphic sites is useful for genetic testing, carrier screening,
CC genotyping or genetic profiling, and identity testing. This
CC polynucleotide is the human cystic fibrosis conductance transmembrane
CC regulator (CFTR) exon 10 DNA sequence containing single nucleotide
CC polymorphisms, used in an exemplification of the invention.
XX

SQ Sequence 831 BP; 263 A; 140 C; 141 G; 287 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 9; Length 831;
Best Local Similarity 63.6%; Pred. No. 0.44;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
        ||| ||||| || ||| ||||| || ||| ||||| ||| ||| ||| |||
Db      328 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTCATTCTG 387

Qy      65 TTGTCACTTTCCGAGGA 81
        || ||| ||| | |||
Db      388 TTCTCAGTTTTCCTGGA 404
```

RESULT 21

AAT04005

ID AAT04005 standard; cDNA; 2640 BP.

XX

AC AAT04005;

XX

DT 25-MAR-2003 (revised)

DT 02-MAY-1996 (first entry)

XX

DE Truncated cystic fibrosis transmembrane conductance regulator cDNA.

XX

KW Cystic fibrosis; transmembrane conductance; N-terminal; soluble;

KW truncated; chloride ion channel; gene therapy; CFTR; regulator;

KW epithelial cells; anion; recombinant production; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 133..2640

FT /*tag= a

FT /note= "truncated N-terminal CFTR protein"

XX

PN WO9525796-A1.

XX

PD 28-SEP-1995.

XX

PF 23-MAR-1995; 95WO-US003680.

XX

PR 23-MAR-1994; 94US-00216971.

XX

PA (IOWA) UNIV IOWA STATE RES FOUND INC.

XX

PI Welsh MJ, Sheppard DM;

XX

DR WPI; 1995-344617/44.

DR P-PSDB; AAR79835.

XX

PT New truncated CFTR polypeptide - functions as a regulated epithelial cell

PT anion channel, used for treating cystic fibrosis.

XX

PS Claim 5; Page 67-70; 85pp; English.

XX

CC AAT04005 encodes AAR79835 a truncated N-terminal portion of the cystic
CC fibrosis transmembrane conductance regulator (CFTR), which can be used to
CC regulate the opening and closing of epithelial cell anion (chloride ion)
CC channels. The truncated cDNA is useful in CF gene therapy, as it is more
CC readily accommodated by available gene therapy vectors, and more easily
CC expressed than full length CFTR. The expressed truncated CFTR protein may
CC be more soluble and therefore more readily purified from host cells,
CC useful in the recombinant prodn. of CFTR. (Updated on 25-MAR-2003 to
CC correct PI field.)

XX

SQ Sequence 2640 BP; 836 A; 509 C; 584 G; 711 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 2; Length 2640;
Best Local Similarity 63.6%; Pred. No. 0.64;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| ||||| || | | | ||| || | ||| | || |||
Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 1604

Qy 65 TTGTCACCTTCCGAGGA 81
|| ||| ||| | |||
Db 1605 TTCTCAGTTTCCTGGA 1621

RESULT 22

ABQ73521

ID ABQ73521 standard; DNA; 3069 BP.

XX

AC ABQ73521;

XX

DT 02-OCT-2002 (first entry)

XX

DE Mouse factor VIII PTM nucleotide sequence.

XX

KW Pre-trans-splicing molecule; PTM; spliceosome; cytostatic; gene therapy;
KW immunosuppressive; antimicrobial; gene regulation; gene repair; cancer;
KW targeted cell death; genetic disorder; infectious disorder;
KW autoimmune disease; proliferative disorder; gene; ds.

XX

OS Mus sp.

OS Synthetic.

XX

PN WO200253581-A2.

XX

PD 11-JUL-2002.

XX

PF 08-JAN-2002; 2002WO-US000416.

XX

PR 08-JAN-2001; 2001US-00756095.

PR 08-JAN-2001; 2001US-00756096.

PR 08-JAN-2001; 2001US-00756097.

PR 20-APR-2001; 2001US-00838858.

PR 29-AUG-2001; 2001US-00941492.

XX

PA (INTR-) INTRONN INC.

XX

PI Mitchell LG, Garcia-Blanco MA, Baker CC, Puttaraju M;
PI Mansfield GS, Chao H;

XX

DR WPI; 2002-566693/60.

XX

PT Novel cell having pre-trans-splicing molecules with target binding
PT domains that target binding of PTM to pre-mRNA, 3' or 5' splice region,
PT spacer region, nucleotide sequence to be trans-spliced to target-pre-
PT mRNA.

XX

PS Example; Fig 43B; 229pp; English.

XX

CC The present invention describes a cell (I) comprising pre-trans-splicing
CC molecules (PTMs) (II) which have one or more target binding domains (IIa)
CC that target binding of PTM to pre-mRNA, 3' splice region (IIb) that
CC includes branch point pyrimidine tract and 3'splice acceptor site, or 5'
CC splice site (IIc), spacer region (IId) that separates RNA splice site
CC from target binding domain, and nucleotide sequence to (IIe) be trans-
CC spliced to target-pre-mRNA. Optionally, the cell comprises (II) either
CC comprising: (A) (IIb) and (IIe); or (B) (IIc), (IId) and (IIe). The cell
CC may comprise a recombinant vector expressing (II). (I) has cytostatic,
CC immunosuppressive and antimicrobial activities, and can be used in gene
CC therapy. (II) comprising one or more (preferably two or more) (IIa) and
CC (IIb) (or (IIc)), (IId) and (IIe), or (II) comprising either (A) or (B)
CC (excluding (IId)), is useful for producing a chimeric RNA molecule in a
CC cell which involves contacting a target pre-mRNA expressed in the cell
CC with (II) that is recognised by nuclear splicing components. The chimeric
CC RNA produced comprises sequences encoding a toxin or translatable
CC protein. The nucleotide sequence to be trans-spliced to target pre-mRNA
CC preferably comprises nucleotide sequences comprising exons 1-10 of cystic
CC fibrosis trans-membrane conductance regulator (CFTR). The chimeric RNA
CC molecule produced using (II) which either comprises (A) or (B) further
CC comprises a nucleotide sequence tag. (I) can be used for gene regulation,
CC gene repair and targeted cell death. (I) can be used for the treatment of
CC various diseases including genetic, infectious or autoimmune diseases and
CC proliferative disorders such as cancer and to regulate gene expression in
CC plants. ABQ73414 to ABQ73536 represent sequences used in the
CC exemplification of the present invention

XX

SQ Sequence 3069 BP; 955 A; 609 C; 662 G; 843 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 6; Length 3069;

Best Local Similarity 63.6%; Pred. No. 0.67;

Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| ||||| || | | | ||| || | ||| | ||| |||

Db 21 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 80

Qy 65 TTGTCACTTTCCGAGGA 81

|| ||| ||| | |||

Db 81 TTCTCAGTTTCCTGGA 97

RESULT 23

AAF84742

ID AAF84742 standard; DNA; 4443 BP.

XX
 AC AAF84742;
 XX
 DT 29-JUN-2001 (first entry)
 XX
 DE DNA encoding cystic fibrosis transmembrane conductance regulator (CFTR).
 XX
 KW Cystic fibrosis transmembrane conductance regulator; CFTR;
 KW cystic fibrosis; CFTR dimerisation; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..4443
 FT /*tag= a
 FT /transl_except= (pos: 2497..2499, aa: Leu)
 FT /product= "cystic fibrosis transmembrane conductance
 FT regulator (CFTR)"
 XX
 PN WO200125421-A2.
 XX
 PD 12-APR-2001.
 XX
 PF 06-OCT-2000; 2000WO-US027900.
 XX
 PR 06-OCT-1999; 99US-0157996P.
 PR 11-FEB-2000; 2000US-0181892P.
 PR 14-FEB-2000; 2000US-0182373P.
 XX
 PA (UYFL) UNIV FLORIDA STATE RES FOUND INC.
 XX
 PI Teem JL;
 XX
 DR WPI; 2001-273576/28.
 DR P-PSDB; AAB68049.
 XX
 PT Detecting interaction of cystic fibrosis transmembrane conductance
 PT regulator (CFTR) polypeptides, useful for screening compounds for
 PT treating cystic fibrosis, comprises using yeast dual hybrid assay.
 XX
 PS Disclosure; Page 41-45; 52pp; English.
 XX
 CC The present sequence encodes a human cystic fibrosis transmembrane
 CC conductance regulator (CFTR) polypeptide. The specification describes a
 CC method for detecting or determining the interaction of two CFTR
 CC polypeptides. The method comprises contacting the CFTR polypeptides and
 CC determining whether the polypeptides interact, where if interaction
 CC occurs a detectable signal or change is induced in the assay system.
 CC Polypeptides and polynucleotides that facilitate the interaction of CFTR
 CC polypeptides are useful for treating cystic fibrosis. Host cells
 CC comprising the CFTR polynucleotide can be used to model wild-type CFTR
 CC protein dimerisation, the effect of cystic fibrosis mutations on
 CC dimerisation and to determine whether a particular mutation of one or
 CC both the CFTR proteins will effect dimerisation of the CFTR proteins and
 CC screen for drugs or compounds that can restore or enhance dimerisation of
 CC CFTR proteins that contain mutations impacting dimerisation
 XX

SQ Sequence 4443 BP; 1363 A; 873 C; 971 G; 1236 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 4; Length 4443;
Best Local Similarity 63.6%; Pred. No. 0.75;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
          || | ||| ||   || | ||| ||   | | | ||| || | ||| | || |||
Db      1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTCATTCTG 1472

Qy      65 TTGTCACTTTCCGAGGA 81
          || ||| ||| | |||
Db      1473 TTCTCAGTTTTCCTGGA 1489
```

RESULT 24

ABX16100

ID ABX16100 standard; cDNA; 4443 BP.

XX

AC ABX16100;

XX

DT 08-APR-2003 (first entry)

XX

DE Human cDNA encoding CFTR mutant I539T/R553M/R555K.

XX

KW Human; ss; gene; CFTR; cystic fibrosis; mutant; CFTR chloride channel;

KW cystic fibrosis transmembrane conductance regulator; gene therapy;

KW cystic fibrosis; I539T/R553M/R555K.

XX

OS Homo sapiens.

OS Synthetic.

XX

FH Key Location/Qualifiers

FT CDS 1. .4443

FT /*tag= a

FT /product= "CFTR I539T/R553M/R555K"

FT /transl_except= (pos:2496. .2499,aa:Leu)

FT mutation replace(1616,T)

FT /*tag= b

FT mutation replace(1656. .1659,CGA)

FT /*tag= c

FT mutation replace(1664,G)

FT /*tag= d

XX

PN US6468793-B1.

XX

PD 22-OCT-2002.

XX

PF 22-OCT-1999; 99US-00425453.

XX

PR 23-OCT-1998; 98US-0105444P.

XX

PA (UYFL) UNIV FLORIDA STATE RES FOUND INC.

XX

PI Teem JL;

XX

DR WPI; 2003-182092/18.

DR P-PSDB; ABG74141.
 XX
 PT Novel cystic fibrosis transmembrane conductance regulator polynucleotide
 PT useful for treating cystic fibrosis, encodes cystic fibrosis
 PT transmembrane conductance regulator polypeptide.
 XX
 PS Claim 4; Col 79-84; 66pp; English.
 XX
 CC The invention relates to a modified cystic fibrosis transmembrane
 CC conductance regulator (CFTR) polynucleotide encoding a CFTR polypeptide,
 CC or its biologically active fragment, where expression of the modified
 CC CFTR protein within a cell results in increased CFTR chloride channel
 CC activity as compared to wild-type CFTR protein. Also included are an
 CC isolated cell comprising the CFTR polynucleotide and a polynucleotide
 CC expression vector comprising the CFTR polynucleotide. The CFTR
 CC polynucleotide is useful for treating cystic fibrosis by gene therapy and
 CC for increasing CFTR-mediated chloride channel activity in a cell. The
 CC CFTR polynucleotide is also useful for treating a patient having a
 CC deficiency or dysfunction in CFTR function. The present sequence encodes
 CC a modified CFTR where the modification comprises Ile at position 539
 CC changed to Thr, Arg at 553 to Met and Arg at 555 Lys
 XX
 SQ Sequence 4443 BP; 1364 A; 873 C; 970 G; 1236 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 8; Length 4443;
 Best Local Similarity 63.6%; Pred. No. 0.75;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || ||| |||| || | | | || | | ||| | || ||
 Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG 1472
 Qy 65 TTGTCACTTCCGAGGA 81
 || ||| ||| | |||
 Db 1473 TTCTCAGTTTTCCTGGA 1489

RESULT 25

ABX16094

ID ABX16094 standard; cDNA; 4443 BP.

XX

AC ABX16094;

XX

DT 08-APR-2003 (first entry)

XX

DE Human cDNA encoding CFTR mutant I539T.

XX

KW Human; ss; gene; CFTR; cystic fibrosis; mutant; CFTR chloride channel;

KW cystic fibrosis transmembrane conductance regulator; gene therapy;

KW cystic fibrosis; I539T.

XX

OS Homo sapiens.

OS Synthetic.

XX

FH Key Location/Qualifiers

FT CDS 1..4443

FT /*tag= a

FT /product= "CFTR I539T"
 FT /transl_except= (pos:2496. .2499,aa:Leu)
 FT mutation replace(1616,T)
 FT /*tag= b
 XX
 PN US6468793-B1.
 XX
 PD 22-OCT-2002.
 XX
 PF 22-OCT-1999; 99US-00425453.
 XX
 PR 23-OCT-1998; 98US-0105444P.
 XX
 PA (UYFL) UNIV FLORIDA STATE RES FOUND INC.
 XX
 PI Teem JL;
 XX
 DR WPI; 2003-182092/18.
 DR P-PSDB; ABG74135.
 XX
 PT Novel cystic fibrosis transmembrane conductance regulator polynucleotide
 PT useful for treating cystic fibrosis, encodes cystic fibrosis
 PT transmembrane conductance regulator polypeptide.
 XX
 PS Claim 2; Col 11-16; 66pp; English.
 XX
 CC The invention relates to a modified cystic fibrosis transmembrane
 CC conductance regulator (CFTR) polynucleotide encoding a CFTR polypeptide,
 CC or its biologically active fragment, where expression of the modified
 CC CFTR protein within a cell results in increased CFTR chloride channel
 CC activity as compared to wild-type CFTR protein. Also included are an
 CC isolated cell comprising the CFTR polynucleotide and a polynucleotide
 CC expression vector comprising the CFTR polynucleotide. The CFTR
 CC polynucleotide is useful for treating cystic fibrosis by gene therapy and
 CC for increasing CFTR-mediated chloride channel activity in a cell. The
 CC CFTR polynucleotide is also useful for treating a patient having a
 CC deficiency or dysfunction in CFTR function. The present sequence encodes
 CC a modified CFTR where the modification comprises Ile at position 539
 CC changed to Thr
 XX
 SQ Sequence 4443 BP; 1363 A; 874 C; 971 G; 1235 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 8; Length 4443;
 Best Local Similarity 63.6%; Pred. No. 0.75;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || ||| |||| || | | | || | || | ||
 Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTTCATTCTG 1472
 Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 1473 TTCTCAGTTTTCTGGA 1489

RESULT 26
 ABX16099

ID ABX16099 standard; cDNA; 4443 BP.
XX
AC ABX16099;
XX
DT 08-APR-2003 (first entry)
XX
DE Human cDNA encoding CFTR mutant I539T/G550E.
XX
KW Human; ss; gene; CFTR; cystic fibrosis; mutant; CFTR chloride channel;
KW cystic fibrosis transmembrane conductance regulator; gene therapy;
KW cystic fibrosis; I539T/G550E.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1. .4443
FT /*tag= a
FT /product= "CFTR I539T/G550E"
FT /transl_except= (pos:2496. .2499,aa:Leu)
FT mutation replace(1616,T)
FT /*tag= b
FT mutation replace(1649,A)
FT /*tag= c
XX
PN US6468793-B1.
XX
PD 22-OCT-2002.
XX
PF 22-OCT-1999; 99US-00425453.
XX
PR 23-OCT-1998; 98US-0105444P.
XX
PA (UYFL) UNIV FLORIDA STATE RES FOUND INC.
XX
PI Teem JL;
XX
DR WPI; 2003-182092/18.
DR P-PSDB; ABG74140.
XX
PT Novel cystic fibrosis transmembrane conductance regulator polynucleotide
PT useful for treating cystic fibrosis, encodes cystic fibrosis
PT transmembrane conductance regulator polypeptide.
XX
PS Claim 3; Col 69-72; 66pp; English.
XX
CC The invention relates to a modified cystic fibrosis transmembrane
CC conductance regulator (CFTR) polynucleotide encoding a CFTR polypeptide,
CC or its biologically active fragment, where expression of the modified
CC CFTR protein within a cell results in increased CFTR chloride channel
CC activity as compared to wild-type CFTR protein. Also included are an
CC isolated cell comprising the CFTR polynucleotide and a polynucleotide
CC expression vector comprising the CFTR polynucleotide. The CFTR
CC polynucleotide is useful for treating cystic fibrosis by gene therapy and
CC for increasing CFTR-mediated chloride channel activity in a cell. The
CC CFTR polynucleotide is also useful for treating a patient having a
CC deficiency or dysfunction in CFTR function. The present sequence encodes

CC a modified CFTR where the modification comprises Ile at position 539
CC changed to Thr and Gly at 550 to Glu
XX
SQ Sequence 4443 BP; 1364 A; 874 C; 970 G; 1235 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 8; Length 4443;
Best Local Similarity 63.6%; Pred. No. 0.75;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
||| ||||| || ||||| || ||| ||| ||| ||| ||| |||
Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTTTCATTCTG 1472

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 1473 TTCTCAGTTTTCCTGGA 1489

RESULT 27

ABX16097

ID ABX16097 standard; cDNA; 4443 BP.

XX

AC ABX16097;

XX

DT 08-APR-2003 (first entry)

XX

DE Human cDNA encoding CFTR mutant R553M.

XX

KW Human; ss; gene; CFTR; cystic fibrosis; mutant; CFTR chloride channel;

KW cystic fibrosis transmembrane conductance regulator; gene therapy;

KW cystic fibrosis; R553M.

XX

OS Homo sapiens.

OS Synthetic.

XX

FH Key Location/Qualifiers

FT CDS 1. .4443

FT /*tag= a

FT /product= "CFTR R553M"

FT /transl_except= (pos:2496. .2499,aa:Leu)

FT mutation replace(1656. .1659,AGA)

FT /*tag= b

XX

PN US6468793-B1.

XX

PD 22-OCT-2002.

XX

PF 22-OCT-1999; 99US-00425453.

XX

PR 23-OCT-1998; 98US-0105444P.

XX

PA (UYFL) UNIV FLORIDA STATE RES FOUND INC.

XX

PI Teem JL;

XX

DR WPI; 2003-182092/18.

DR P-PSDB; ABG74138.

XX
PT Novel cystic fibrosis transmembrane conductance regulator polynucleotide
PT useful for treating cystic fibrosis, encodes cystic fibrosis
PT transmembrane conductance regulator polypeptide.
XX
PS Example 2; Col 45-50; 66pp; English.
XX
CC The invention relates to a modified cystic fibrosis transmembrane
CC conductance regulator (CFTR) polynucleotide encoding a CFTR polypeptide,
CC or its biologically active fragment, where expression of the modified
CC CFTR protein within a cell results in increased CFTR chloride channel
CC activity as compared to wild-type CFTR protein. Also included are an
CC isolated cell comprising the CFTR polynucleotide and a polynucleotide
CC expression vector comprising the CFTR polynucleotide. The CFTR
CC polynucleotide is useful for treating cystic fibrosis by gene therapy and
CC for increasing CFTR-mediated chloride channel activity in a cell. The
CC CFTR polynucleotide is also useful for treating a patient having a
CC deficiency or dysfunction in CFTR function. The present sequence encodes
CC a modified CFTR where the modification comprises Arg at position 553
CC changed to Met
XX
SQ Sequence 4443 BP; 1363 A; 872 C; 971 G; 1237 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 8; Length 4443;
Best Local Similarity 63.6%; Pred. No. 0.75;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| ||||| || | | ||| || | ||| | || |||
Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTTCATTCTG 1472
Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 1473 TTCTCAGTTTTCCTGGA 1489

RESULT 28

ABX16103

ID ABX16103 standard; cDNA; 4443 BP.

XX

AC ABX16103;

XX

DT 08-APR-2003 (first entry)

XX

DE Human cDNA encoding CFTR mutant I539M/G550E.

XX

KW Human; ss; gene; CFTR; cystic fibrosis; mutant; CFTR chloride channel;

KW cystic fibrosis transmembrane conductance regulator; gene therapy;

KW cystic fibrosis; I539T/G550E.

XX

OS Homo sapiens.

OS Synthetic.

XX

FH Key Location/Qualifiers

FT CDS 1..4443

FT /*tag= a

FT /product= "CFTR I539T/G550E"

```

FT          /transl_except= (pos:2496. .2499,aa:Leu)
FT mutation replace(1617,A)
FT          /*tag= b
FT mutation replace(1649,A)
FT          /*tag= c
XX
PN US6468793-B1.
XX
PD 22-OCT-2002.
XX
PF 22-OCT-1999; 99US-00425453.
XX
PR 23-OCT-1998; 98US-0105444P.
XX
PA (UYFL ) UNIV FLORIDA STATE RES FOUND INC.
XX
PI Teem JL;
XX
DR WPI; 2003-182092/18.
DR P-PSDB; ABG74144.
XX
PT Novel cystic fibrosis transmembrane conductance regulator polynucleotide
PT useful for treating cystic fibrosis, encodes cystic fibrosis
PT transmembrane conductance regulator polypeptide.
XX
PS Disclosure; Col 113-118; 66pp; English.
XX
CC The invention relates to a modified cystic fibrosis transmembrane
CC conductance regulator (CFTR) polynucleotide encoding a CFTR polypeptide,
CC or its biologically active fragment, where expression of the modified
CC CFTR protein within a cell results in increased CFTR chloride channel
CC activity as compared to wild-type CFTR protein. Also included are an
CC isolated cell comprising the CFTR polynucleotide and a polynucleotide
CC expression vector comprising the CFTR polynucleotide. The CFTR
CC polynucleotide is useful for treating cystic fibrosis by gene therapy and
CC for increasing CFTR-mediated chloride channel activity in a cell. The
CC CFTR polynucleotide is also useful for treating a patient having a
CC deficiency or dysfunction in CFTR function. The present sequence encodes
CC a modified CFTR where the modification comprises Ile at position 539
CC changed to Met and Gly at 550 to Glu
XX
SQ Sequence 4443 BP; 1363 A; 873 C; 971 G; 1236 T; 0 U; 0 Other;

```

```

Query Match          31.6%; Score 32.2; DB 8; Length 4443;
Best Local Similarity 63.6%; Pred. No. 0.75;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy          5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
              || | |||| ||      ||| ||||| ||  | | | ||| || | ||| | || |||
Db          1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 1472

Qy          65 TTGTCACTTTCCGAGGA 81
              || ||| ||| | |||
Db          1473 TTCTCAGTTTTCCTGGA 1489

```

RESULT 29

ABX16095

ID ABX16095 standard; cDNA; 4443 BP.

XX

AC ABX16095;

XX

DT 08-APR-2003 (first entry)

XX

DE Human cDNA encoding CFTR mutant I539M.

XX

KW Human; ss; gene; CFTR; cystic fibrosis; mutant; CFTR chloride channel;
KW cystic fibrosis transmembrane conductance regulator; gene therapy;
KW cystic fibrosis; I539M.

XX

OS Homo sapiens.

OS Synthetic.

XX

FH Key Location/Qualifiers

FT CDS 1. .4443

FT /*tag= a

FT /product= "CFTR I539M"

FT /transl_except= (pos:2496. .2499,aa:Leu)

FT mutation replace(1617,A)

FT /*tag= b

XX

PN US6468793-B1.

XX

PD 22-OCT-2002.

XX

PF 22-OCT-1999; 99US-00425453.

XX

PR 23-OCT-1998; 98US-0105444P.

XX

PA (UYFL) UNIV FLORIDA STATE RES FOUND INC.

XX

PI Teem JL;

XX

DR WPI; 2003-182092/18.

DR P-PSDB; ABG74136.

XX

PT Novel cystic fibrosis transmembrane conductance regulator polynucleotide
PT useful for treating cystic fibrosis, encodes cystic fibrosis
PT transmembrane conductance regulator polypeptide.

XX

PS Example 1; Col 23-28; 66pp; English.

XX

CC The invention relates to a modified cystic fibrosis transmembrane
CC conductance regulator (CFTR) polynucleotide encoding a CFTR polypeptide,
CC or its biologically active fragment, where expression of the modified
CC CFTR protein within a cell results in increased CFTR chloride channel
CC activity as compared to wild-type CFTR protein. Also included are an
CC isolated cell comprising the CFTR polynucleotide and a polynucleotide
CC expression vector comprising the CFTR polynucleotide. The CFTR
CC polynucleotide is useful for treating cystic fibrosis by gene therapy and
CC for increasing CFTR-mediated chloride channel activity in a cell. The
CC CFTR polynucleotide is also useful for treating a patient having a
CC deficiency or dysfunction in CFTR function. The present sequence encodes
CC a modified CFTR where the modification comprises Ile at position 539

CC changed to Met

XX

SQ Sequence 4443 BP; 1362 A; 873 C; 972 G; 1236 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 8; Length 4443;

Best Local Similarity 63.6%; Pred. No. 0.75;

Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
          || | ||| ||   || | ||| ||   | | | ||| || | ||| | || |||
Db      1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTTCATTCTG 1472

Qy      65 TTGTCACTTTCCGAGGA 81
          || ||| ||| | |||
Db      1473 TTCTCAGTTTTCCTGGA 1489
```

RESULT 30

ABX16098

ID ABX16098 standard; cDNA; 4443 BP.

XX

AC ABX16098;

XX

DT 08-APR-2003 (first entry)

XX

DE Human cDNA encoding CFTR mutant R555K.

XX

KW Human; ss; gene; CFTR; cystic fibrosis; mutant; CFTR chloride channel;

KW cystic fibrosis transmembrane conductance regulator; gene therapy;

KW cystic fibrosis; R555K.

XX

OS Homo sapiens.

OS Synthetic.

XX

FH Key Location/Qualifiers

FT CDS 1..4443

FT /*tag= a

FT /product= "CFTR R555K"

FT /transl_except= (pos:2496..2499,aa:Leu)

FT mutation replace(1664,G)

FT /*tag= b

XX

PN US6468793-B1.

XX

PD 22-OCT-2002.

XX

PF 22-OCT-1999; 99US-00425453.

XX

PR 23-OCT-1998; 98US-0105444P.

XX

PA (UYFL) UNIV FLORIDA STATE RES FOUND INC.

XX

PI Teem JL;

XX

DR WPI; 2003-182092/18.

DR P-PSDB; ABG74139.

XX

PT Novel cystic fibrosis transmembrane conductance regulator polynucleotide
PT useful for treating cystic fibrosis, encodes cystic fibrosis
PT transmembrane conductance regulator polypeptide.

XX

PS Example 2; Col 57-62; 66pp; English.

XX

CC The invention relates to a modified cystic fibrosis transmembrane
CC conductance regulator (CFTR) polynucleotide encoding a CFTR polypeptide,
CC or its biologically active fragment, where expression of the modified
CC CFTR protein within a cell results in increased CFTR chloride channel
CC activity as compared to wild-type CFTR protein. Also included are an
CC isolated cell comprising the CFTR polynucleotide and a polynucleotide
CC expression vector comprising the CFTR polynucleotide. The CFTR
CC polynucleotide is useful for treating cystic fibrosis by gene therapy and
CC for increasing CFTR-mediated chloride channel activity in a cell. The
CC CFTR polynucleotide is also useful for treating a patient having a
CC deficiency or dysfunction in CFTR function. The present sequence encodes
CC a modified CFTR where the modification comprises Arg at position 555
CC changed to Lys

XX

SQ Sequence 4443 BP; 1364 A; 873 C; 970 G; 1236 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 8; Length 4443;
Best Local Similarity 63.6%; Pred. No. 0.75;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| ||||| || | | | || | || | || | ||
Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG 1472

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 1473 TTCTCAGTTTTCCTGGA 1489

RESULT 31

ABX16102

ID ABX16102 standard; cDNA; 4443 BP.

XX

AC ABX16102;

XX

DT 08-APR-2003 (first entry)

XX

DE Human cDNA encoding CFTR mutant I539M/R553M/R555K.

XX

KW Human; ss; gene; CFTR; cystic fibrosis; mutant; CFTR chloride channel;
KW cystic fibrosis transmembrane conductance regulator; gene therapy;
KW cystic fibrosis; I539M/R553M/R555K.

XX

OS Homo sapiens.

OS Synthetic.

XX

FH Key Location/Qualifiers

FT CDS 1. .4443

FT /*tag= a

FT /product= "CFTR I539M/R553M/R555K"

FT /transl_except= (pos:2496. .2499,aa:Leu)

```

FT      mutation      replace(1617,A)
FT      /*tag=  b
FT      mutation      replace(1656. .1659,CGA)
FT      /*tag=  c
FT      mutation      replace(1664,G)
FT      /*tag=  d
XX
PN      US6468793-B1.
XX
PD      22-OCT-2002.
XX
PF      22-OCT-1999;   99US-00425453.
XX
PR      23-OCT-1998;   98US-0105444P.
XX
PA      (UYFL ) UNIV FLORIDA STATE RES FOUND INC.
XX
PI      Teem JL;
XX
DR      WPI; 2003-182092/18.
DR      P-PSDB; ABG74143.
XX
PT      Novel cystic fibrosis transmembrane conductance regulator polynucleotide
PT      useful for treating cystic fibrosis, encodes cystic fibrosis
PT      transmembrane conductance regulator polypeptide.
XX
PS      Disclosure; Col 103-106; 66pp; English.
XX
CC      The invention relates to a modified cystic fibrosis transmembrane
CC      conductance regulator (CFTR) polynucleotide encoding a CFTR polypeptide,
CC      or its biologically active fragment, where expression of the modified
CC      CFTR protein within a cell results in increased CFTR chloride channel
CC      activity as compared to wild-type CFTR protein. Also included are an
CC      isolated cell comprising the CFTR polynucleotide and a polynucleotide
CC      expression vector comprising the CFTR polynucleotide. The CFTR
CC      polynucleotide is useful for treating cystic fibrosis by gene therapy and
CC      for increasing CFTR-mediated chloride channel activity in a cell. The
CC      CFTR polynucleotide is also useful for treating a patient having a
CC      deficiency or dysfunction in CFTR function. The present sequence encodes
CC      a modified CFTR where the modification comprises Ile at position 539
CC      changed to Met, Arg at 553 to Met and Arg at 555 Lys
XX
SQ      Sequence 4443 BP; 1363 A; 872 C; 971 G; 1237 T; 0 U; 0 Other;

```

```

Query Match          31.6%; Score 32.2; DB 8; Length 4443;
Best Local Similarity 63.6%; Pred. No. 0.75;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
        || | |||| ||      ||| ||||| ||      | | | ||| || | ||| | || |||
Db      1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTTTCATTCTG 1472

Qy      65 TTGTCACTTTCCGAGGA 81
        || ||| ||| | |||
Db      1473 TTCTCAGTTTTCCTGGA 1489

```

RESULT 32

ABX16096

ID ABX16096 standard; cDNA; 4443 BP.

XX

AC ABX16096;

XX

DT 08-APR-2003 (first entry)

XX

DE Human cDNA encoding CFTR mutant G550E.

XX

KW Human; ss; gene; CFTR; cystic fibrosis; mutant; CFTR chloride channel;

KW cystic fibrosis transmembrane conductance regulator; gene therapy;

KW cystic fibrosis; G550E.

XX

OS Homo sapiens.

OS Synthetic.

XX

FH Key Location/Qualifiers

FT CDS 1. .4443

FT /*tag= a

FT /product= "CFTR G550E"

FT /transl_except= (pos:2496. .2499,aa:Leu)

FT mutation replace(1649,A)

FT /*tag= b

XX

PN US6468793-B1.

XX

PD 22-OCT-2002.

XX

PF 22-OCT-1999; 99US-00425453.

XX

PR 23-OCT-1998; 98US-0105444P.

XX

PA (UYFL) UNIV FLORIDA STATE RES FOUND INC.

XX

PI Teem JL;

XX

DR WPI; 2003-182092/18.

DR P-PSDB; ABG74137.

XX

PT Novel cystic fibrosis transmembrane conductance regulator polynucleotide

PT useful for treating cystic fibrosis, encodes cystic fibrosis

PT transmembrane conductance regulator polypeptide.

XX

PS Example 2; Col 35-38; 66pp; English.

XX

CC The invention relates to a modified cystic fibrosis transmembrane

CC conductance regulator (CFTR) polynucleotide encoding a CFTR polypeptide,

CC or its biologically active fragment, where expression of the modified

CC CFTR protein within a cell results in increased CFTR chloride channel

CC activity as compared to wild-type CFTR protein. Also included are an

CC isolated cell comprising the CFTR polynucleotide and a polynucleotide

CC expression vector comprising the CFTR polynucleotide. The CFTR

CC polynucleotide is useful for treating cystic fibrosis by gene therapy and

CC for increasing CFTR-mediated chloride channel activity in a cell. The

CC CFTR polynucleotide is also useful for treating a patient having a

CC deficiency or dysfunction in CFTR function. The present sequence encodes

CC a modified CFTR where the modification comprises Gly at position 539
CC changed to Glu

XX

SQ Sequence 4443 BP; 1364 A; 873 C; 970 G; 1236 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 8; Length 4443;

Best Local Similarity 63.6%; Pred. No. 0.75;

Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64

|| | |||| || ||| ||||| || | | | ||| || | ||| | || |||

Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTTTCATTCTG 1472

Qy 65 TTGTCACTTTCCGAGGA 81

|| ||| ||| | |||

Db 1473 TTCTCAGTTTTCCTGGA 1489

RESULT 33

ABX16101

ID ABX16101 standard; cDNA; 4443 BP.

XX

AC ABX16101;

XX

DT 08-APR-2003 (first entry)

XX

DE Human cDNA encoding wild-type CFTR.

XX

KW Human; ss; gene; CFTR; cystic fibrosis; CFTR chloride channel;

KW cystic fibrosis transmembrane conductance regulator; gene therapy;

KW cystic fibrosis.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1. .4443

FT /*tag= a

FT /product= "CFTR"

FT /transl_except= (pos:2496. .2499,aa:Leu)

XX

PN US6468793-B1.

XX

PD 22-OCT-2002.

XX

PF 22-OCT-1999; 99US-00425453.

XX

PR 23-OCT-1998; 98US-0105444P.

XX

PA (UYFL) UNIV FLORIDA STATE RES FOUND INC.

XX

PI Teem JL;

XX

DR WPI; 2003-182092/18.

DR P-PSDB; ABG74142.

XX

PT Novel cystic fibrosis transmembrane conductance regulator polynucleotide

PT useful for treating cystic fibrosis, encodes cystic fibrosis

PT transmembrane conductance regulator polypeptide.
 XX
 PS Example 1; Col 91-96; 66pp; English.
 XX
 CC The invention relates to a modified cystic fibrosis transmembrane
 CC conductance regulator (CFTR) polynucleotide encoding a CFTR polypeptide,
 CC or its biologically active fragment, where expression of the modified
 CC CFTR protein within a cell results in increased CFTR chloride channel
 CC activity as compared to wild-type CFTR protein. Also included are an
 CC isolated cell comprising the CFTR polynucleotide and a polynucleotide
 CC expression vector comprising the CFTR polynucleotide. The CFTR
 CC polynucleotide is useful for treating cystic fibrosis by gene therapy and
 CC for increasing CFTR-mediated chloride channel activity in a cell. The
 CC CFTR polynucleotide is also useful for treating a patient having a
 CC deficiency or dysfunction in CFTR function. The present sequence encodes
 CC wild-type CFTR
 XX
 SQ Sequence 4443 BP; 1363 A; 873 C; 971 G; 1236 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 8; Length 4443;
 Best Local Similarity 63.6%; Pred. No. 0.75;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || ||| |||| || | | ||| || | ||| | || |||
 Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGCTAAAATTAAGCACAGTGGAAGAATTCATTCTG 1472
 Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 1473 TTCTCAGTTTTTCCTGGA 1489

RESULT 34

AAZ11643

ID AAZ11643 standard; cDNA; 4560 BP.

XX

AC AAZ11643;

XX

DT 26-MAY-2000 (first entry)

XX

DE CFTR protein encoding cDNA.

XX

KW AAV vector; inverted terminal repeat; ITR; gene therapy; CFTR; TK gene;

KW cystic fibrosis transmembrane conductance regulator; cystic fibrosis;

KW promoter; HSV; thymidine kinase; chromosome 7q31; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 133..4560

FT /*tag= a

FT /transl_except= (pos: 3580..3582, aa: Ile)

FT /note= "the stop codon is not indicated"

XX

PN WO9943789-A1.

XX

PD 02-SEP-1999.

XX
 PF 25-FEB-1999; 99WO-US004212.
 XX
 PR 25-FEB-1998; 98US-0075980P.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Dong J, Kan YW;
 XX
 DR WPI; 1999-550866/46.
 DR P-PSDB; AAY33968.
 XX
 PT Efficient AAV vectors useful in gene therapy protocols for the treatment
 PT of cystic fibrosis.
 XX
 PS Example 1; Page 33; 34pp; English.
 XX
 CC The invention provides efficient AAV vectors with improved capacity for
 CC DNA due to the removal of all nucleic acid sequences that are not
 CC essential for replication (to leave just 2 inverted terminal repeat
 CC sequences (ITRs)). The AAV vectors may be used for the delivery of
 CC therapeutic nucleic acids in gene therapy protocols. In particular, they
 CC may be used to deliver cystic fibrosis (CF) transmembrane conductance
 CC regulator (CFTR) polynucleotides to the respiratory tract of CF patients
 CC to rectify mutations in the patients own CFTR genes and restore normal
 CC function to the chloride channel the gene encodes. The AAV vector lacks
 CC all nucleic acids that are not essential for replication, therefore
 CC giving it a greater capacity for exogenous DNA and hence improving the
 CC efficiency with which it transfects cells. The AAV vectors of the
 CC invention can efficiently and persistently transfer CFTR polynucleotides
 CC to the airway epithelium of CF patients without any adverse side effects.
 CC The present sequence represents a cDNA encoding the CFTR protein
 XX
 SQ Sequence 4560 BP; 1397 A; 910 C; 1003 G; 1250 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 2; Length 4560;
 Best Local Similarity 63.6%; Pred. No. 0.76;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 ||| ||||| || ||||| || ||| ||||| ||| ||||| |||||
 Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTTCATTCTG 1604
 Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 1605 TTCTCAGTTTTCTGGA 1621

RESULT 35
 AAS81827
 ID AAS81827 standard; cDNA; 4845 BP.
 XX
 AC AAS81827;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #17631.

XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR P-PSDB; ABG17640.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 1; SEQ ID NO 17631; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 4845 BP; 1806 A; 1007 C; 921 G; 1111 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 5; Length 4845;
 Best Local Similarity 63.6%; Préd. No. 0.77;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || ||| |||| || | | | ||| || | ||| | ||
 Db 3555 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTTCATTCTG 3614

Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 3615 TTCTCAGTTTTTCCTGGA 3631

RESULT 36

AAQ13605

ID AAQ13605 standard; cDNA; 4894 BP.

XX

AC AAQ13605;

XX

DT 25-MAR-2003 (revised)

DT 21-NOV-1991 (first entry)

XX

DE Cystic fibrosis transmembrane conductance regulator gene.

XX

KW CFTR; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 133..4575

FT /*tag= a

XX

PN EP446017-A.

XX

PD 11-SEP-1991.

XX

PF 05-MAR-1991; 91EP-00301819.

XX

PR 05-MAR-1990; 90US-00488307.

PR 27-SEP-1990; 90US-00589295.

PR 15-NOV-1990; 90US-00613592.

XX

PA (GENZ) GENZYME CORP.

XX

PI Gregory RJ, Cheng SH, Smith A, Paul S, Hehir KM, Marshall J;

XX

DR WPI; 1991-268856/37.

DR P-PSDB; AAR13894.

XX

PT DNA encoding cystic fibrosis trans-membrane regulator - for use in

PT treating and diagnosing cystic fibrosis.

XX

PS Claim 1; Page 26; 50pp; English.

XX

CC The DNA sequence codes for cystic fibrosis transmembrane regulator
 CC (CFTR). It may be used in gene therapy to obtain in vivo prodn. of CFTR
 CC in cystic fibrosis patients, and also in the prodn. of CFTR for protein
 CC replacement therapy. CFTR may also be used in the diagnosis of cystic
 CC fibrosis by monitoring its presence or absence. See also AAQ13606.
 CC (Updated on 25-MAR-2003 to correct PA field.)

XX

SQ Sequence 4894 BP; 1495 A; 960 C; 1094 G; 1345 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 2; Length 4894;
Best Local Similarity 63.6%; Pred. No. 0.77;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
        || | ||| ||   ||| |||| ||   | | | ||| || | ||| | || |||
Db      1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 1604

Qy      65 TTGTCACTTTCCGAGGA 81
        || ||| ||| | |||
Db      1605 TTCTCAGTTTTTCCTGGA 1621
```

RESULT 37

AAQ68002

ID AAQ68002 standard; DNA; 5635 BP.

XX

AC AAQ68002;

XX

DT 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 26-OCT-1995 (first entry)

XX

DE Ad2/CFTR-1 nucleotide sequence.

XX

KW Recombinant adenovirus; Ad2/CFTR-1; adenovirus 2 serotype; Ela; Elb;

KW viral replication; gene expression; gene therapy; cystic fibrosis;

KW cystic fibrosis transmembrane conductance regulator; CFTR; promoter; E3;

KW p19; MHC; class 1; viral latency; pulmonary airway; ds.

XX

OS Homo sapiens.

OS Unidentified.

XX

FH Key Location/Qualifiers

FT repeat_region 1. .104

FT /*tag= a

FT /rpt_type= INVERTED

FT /note= "Represents the origin of replication"

FT enhancer 190. .380

FT /*tag= b

FT /function= "E1A enhancer and viral packaging domain"

FT promoter 380. .500

FT /*tag= c

FT /note= "E1A promoter region"

FT prim_transcript 499. .5635

FT /*tag= d

FT /note= "Hybrid E1A-CFTR-E1B message"

FT 5'UTR 499. .546

FT /*tag= e

FT misc_feature 547. .595

FT /*tag= f

FT /note= "Synthetic linker sequences"

FT misc_feature 593. .5093

FT /*tag= g

FT /note= "Represents nucleotides 123-4622 of the published

```

FT          CFTR cDNA sequence"
FT CDS      603. .5045
FT          /*tag= h
FT          /product= "CFTR"
FT 3'UTR    5093. .5635
FT          /*tag= i
FT          /note= "ElB 3' UTR"
FT intron   5099. .5190
FT          /*tag= j
FT          /note= "ElB 3' intron"
FT prim_transcript 5177. .5635
FT          /*tag= k
FT          /note= "IX protein mRNA"
FT CDS      5201. .5623
FT          /*tag= l
FT          /product= "IX protein (Hexon-associated protein)"
XX
PN WO9412649-A2.
XX
PD 09-JUN-1994.
XX
PF 02-DEC-1993; 93WO-US011667.
XX
PR 03-DEC-1992; 92US-00985478.
PR 01-OCT-1993; 93US-00130682.
PR 13-OCT-1993; 93US-00136742.
XX
PA (GENZ ) GENZYME CORP.
XX
PI Gregory RJ, Armentano D, Couture LA, Smith AE;
XX
DR WPI; 1994-200277/24.
DR P-PSDB; AAR79011, AAR79012.
XX
PT Adeno:virus-based gene therapy vectors - esp. useful for gene therapy of
PT cystic fibrosis.
XX
PS Claim 4; Page 67-80; 167pp; English.
XX
CC This sequence represents the nucleotide sequence of the recombinant
CC adenovirus Ad2/CFTR-1. This virus is derived from the relatively benign
CC adenovirus 2 serotype. The Ela and Elb regions of the viral genome, which
CC are involved in the early stages of viral replication have been deleted
CC which impairs viral gene expression and viral replication. The cystic
CC fibrosis transmembrane conductance regulator (CFTR) coding sequence is
CC inserted into the genome in place of the Ela/Elb region and transcription
CC of the CFTR sequence is driven by the endogenous Ela promoter. This is a
CC moderately strong promoter that is functional in a variety of cells. This
CC adenovirus retains the E3 viral coding region. As a consequence the
CC length of the adenovirus-CFTR DNA is greater than that of wild type
CC adenovirus. This renders the DNA more difficult to package and means that
CC the growth of the Ad2/CFTR virus is impaired even in permissive cells
CC that provide the missing Ela and Elb functions. The E3 region encodes a
CC number of proteins, including p19 which is believed to interact with and
CC prevent presentation of MHC class 1 proteins. This property prevents
CC recognition of the infected cells and thus may allow viral latency. This
CC adenovirus may be administered to the pulmonary airways in the gene

```

CC therapy of cystic fibrosis. (Updated on 25-MAR-2003 to correct PN field.)
 CC (Updated on 27-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 5635 BP; 1619 A; 1142 C; 1324 G; 1550 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 2; Length 5635;
 Best Local Similarity 63.6%; Pred. No. 0.81;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || ||| |||| || | | |||| || | ||| | || |||
 Db 2015 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTTCATTCTG 2074
 QY 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 2075 TTCTCAGTTTTCTGGA 2091

RESULT 38

AAQ13053

ID AAQ13053 standard; cDNA; 6126 BP.

XX

AC AAQ13053;

XX

DT 14-OCT-1991 (first entry)

XX

DE CFTR delta I507.

XX

KW Deletion; mutant; diagnosis; antibodies; drug therapy;

KW ATP-binding domain; ss.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	CDS	133..4569
----	-----	-----------

FT		/*tag= a
----	--	----------

FT		/label= CFTR-mutant
----	--	---------------------

FT	misc_feature	185..186
----	--------------	----------

FT		/*tag= b
----	--	----------

FT		/label= exon_junction
----	--	-----------------------

FT	misc_feature	296..297
----	--------------	----------

FT		/*tag= c
----	--	----------

FT		/label= exon_junction
----	--	-----------------------

FT	misc_feature	372..438
----	--------------	----------

FT		/*tag= w
----	--	----------

FT		/label= membrane-spanning_segment
----	--	-----------------------------------

FT	misc_feature	405..406
----	--------------	----------

FT		/*tag= d
----	--	----------

FT		/label= exon_junction
----	--	-----------------------

FT	misc_feature	484..546
----	--------------	----------

FT		/*tag= x
----	--	----------

FT		/label= membrane-spanning_segment
----	--	-----------------------------------

FT	misc_feature	621..622
----	--------------	----------

FT		/*tag= e
----	--	----------

FT		/label= exon_junction
----	--	-----------------------

FT	misc_feature	711..712
----	--------------	----------

FT		/*tag= f
----	--	----------

```

FT                                     /label= exon_junction
FT  misc_feature                     714. .777
FT                                     /*tag= y
FT                                     /label= membrane-spanning_segment
FT  misc_feature                     793. .855
FT                                     /*tag= z
FT                                     /label= membrane-spanning_segment
FT  misc_feature                     1001. .1002
FT                                     /*tag= g
FT                                     /label= exon_junction
FT  misc_feature                     1054. .1116
FT                                     /*tag= aa
FT                                     /label= membrane-spanning_segment
FT  misc_feature                     1120. .1182
FT                                     /*tag= ab
FT                                     /label= membrane-spanning_segment
FT  misc_feature                     1248. .1249
FT                                     /*tag= h
FT                                     /label= exon_junction
FT  misc_feature                     1341. .1342
FT                                     /*tag= i
FT                                     /label= exon_junction
FT  misc_binding                     1429. .1881
FT                                     /*tag= ai
FT                                     /label= ATP-binding_fold
FT  misc_feature                     1523. .1524
FT                                     /*tag= j
FT                                     /label= exon_junction
FT  misc_feature                     1713. .1714
FT                                     /*tag= k
FT                                     /label= exon_junction
FT  misc_feature                     1808. .1809
FT                                     /*tag= l
FT                                     /label= exon_junction
FT  misc_feature                     1895. .1896
FT                                     /*tag= m
FT                                     /label= exon_junction
FT  misc_feature                     2619. .2620
FT                                     /*tag= n
FT                                     /label= exon_junction
FT  misc_feature                     2707. .2769
FT                                     /*tag= ac
FT                                     /label= membrane-spanning_segment
FT  misc_feature                     2786. .2787
FT                                     /*tag= o
FT                                     /label= exon_junction
FT  misc_feature                     2863. .2925
FT                                     /*tag= ad
FT                                     /label= membrane-spanning_segment
FT  misc_feature                     3037. .3038
FT                                     /*tag= p
FT                                     /label= exon_junction
FT  misc_feature                     3100. .3162
FT                                     /*tag= ae
FT                                     /label= membrane-spanning_segment
FT  misc_feature                     3168. .3231
FT                                     /*tag= af

```

FT /label= membrane-spanning_segment
 FT misc_feature 3436. .3498
 FT /*tag= ag
 FT /label= membrane-spanning_segment
 FT misc_feature 3496. .3497
 FT /*tag= q
 FT /label= exon_junction
 FT misc_feature 3514. .3579
 FT /*tag= ah
 FT /label= membrane-spanning_segment
 FT misc_feature 3596. .3597
 FT /*tag= r
 FT /label= exon_junction
 FT misc_binding 3784. .4287
 FT /*tag= aj
 FT /label= ATP-binding_fold
 FT misc_feature 3846. .3847
 FT /*tag= s
 FT /label= exon_junction
 FT misc_feature 4092. .4093
 FT /*tag= t
 FT /label= exon_junction
 FT misc_feature 4265. .4266
 FT /*tag= u
 FT /label= exon_junction
 FT misc_feature 4371. .4372
 FT /*tag= v
 FT /label= exon_junction

XX

PN WO9110734-A.

XX

PD 25-JUL-1991.

XX

PF 12-JAN-1990; 90CA-02007699.

XX

PR 12-JAN-1990; 90CA-02007699.

PR 01-MAR-1990; 90CA-02011253.

PR 10-JUL-1990; 90CA-02020817.

XX

PA (HSCR-) HSC RES DEV CORP.

XX

PI Tsui LC, Rommens JM, Kerem B;

XX

DR WPI; 1991-238022/32.

DR P-PSDB; AAR13231.

XX

PT Mutant cystic fibrosis trans-membrane conductance regulator gene - used
 PT for producing prods. for diagnosis, screening and therapy of cystic
 PT fibrosis.

XX

PS Claim 1; Page 121; 178pp; English.

XX

CC The deletion of the 3 bp (ATC) at the I506 or I507 position results in
 CC the loss of an isoleucine residue from the putative CFTR, within the same
 CC ATP-binding domain where deltaF508 resides, but it is not evident whether
 CC this deleted amino acid corresponds to the position 506 or 507. Since the
 CC 506 and 507 positions are repeats, it is at present impossible to

CC determine in which position the 3 bp deletion occurs. Nucleotide 6126 is
CC followed by a poly(dA) tract. The mutant CF gene when expressed in cells
CC of the human body, is associated with altered cell function which
CC correlates with the genetic disease cystic fibrosis. See also AAQ13053-72
XX
SQ Sequence 6126 BP; 1884 A; 1182 C; 1329 G; 1731 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 2; Length 6126;
Best Local Similarity 63.6%; Pred. No. 0.83;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| |||| || | | | || | | || | || |
Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGGAGAATTCATTCTG 1604

Qy 65 TTGTCACTTTCGAGGA 81
|| ||| ||| | |||
Db 1605 TTCTCAGTTTCCTGGA 1621

RESULT 39

AAX35553

ID AAX35553 standard; DNA; 6126 BP.

XX

AC AAX35553;

XX

DT 08-JUL-1999 (first entry)

XX

DE DeltaF508 cystic fibrosis transmembrane conductance regulator DNA.

XX

KW Flavone; isoflavone; resveratrol; ascorbic acid; ascorbate salt;

KW dehydroascorbic acid; chloride transport; epithelial cell;

KW cystic fibrosis; chloride ion conductance;

KW cystic fibrosis transmembrane conductance regulator; CFTR;

KW chronic bronchitis; asthma; intestinal constipation; ss.

XX

OS Homo sapiens.

XX

PN WO9918953-A1.

XX

PD 22-APR-1999.

XX

PF 16-OCT-1998; 98WO-US021887.

XX

PR 16-OCT-1997; 97US-00951912.

XX

PA (CHIL-) CHILDREN'S HOSPITAL OAKLAND RES INST.

XX

PI Fischer HB, Illek B;

XX

DR WPI; 1999-277427/23.

DR P-PSDB; AAY02279.

XX

PT Use of flavones and isoflavones - for stimulating chloride transport in
PT epithelial cells and treating cystic fibrosis.

XX

PS Disclosure; Page 70-73; 97pp; English.

XX
 CC The specification describes compounds comprising flavones/isoflavones,
 CC resveratrol, ascorbic acid, ascorbate salts and/or dehydroascorbic acid
 CC which can be used for stimulating chloride transport in epithelial cells
 CC and treating cystic fibrosis. The compounds can be used to increase
 CC chloride ion conductance in airway epithelial cells or intestine,
 CC pancreas, gallbladder, sweat duct, salivary gland or mammary epithelial
 CC cells. The compounds are useful for treating a patient with cystic
 CC fibrosis, where the patient's cystic fibrosis transmembrane conductance
 CC regulator (CFTR) protein has a deletion at position 508 or point mutation
 CC at 551. They may also be used for treating chronic bronchitis, asthma and
 CC intestinal constipation. The present sequence encodes a human CFTR
 CC protein with a F508 deletion mutation
 XX
 SQ Sequence 6126 BP; 1886 A; 1181 C; 1330 G; 1729 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 2; Length 6126;
 Best Local Similarity 63.6%; Pred. No. 0.83;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 ||| ||||| || ||| ||||| || | | | ||| || | ||| ||| |||
 Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG 1604
 Qy 65 TTGTCACTTTCGAGGA 81
 || ||| ||| | |||
 Db 1605 TTCTCAGTTTTCCTGGA 1621

RESULT 40

AAS20529

ID AAS20529 standard; DNA; 6126 BP.

XX

AC AAS20529;

XX

DT 23-APR-2002 (first entry)

XX

DE Human delta-F508-CFTR DNA.

XX

KW Human; cystic fibrosis transmembrane conductance regulator; CFTR; gene;
 KW flavone; isoflavone; chloride transport; epithelial tissue; mucus; ds;
 KW cystic fibrosis; chronic bronchitis; asthma; delta-F508-CFTR.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 133..4572

FT /*tag= a

FT /product= "Human delta-F508-CFTR protein"

XX

PN US6329422-B1.

XX

PD 11-DEC-2001.

XX

PF 16-OCT-1998; 98US-00174077.

XX

PR 16-OCT-1997; 97US-00951912.

XX
PA (CHIL-) CHILDREN'S HOSPITAL OAKLAND RES INST.
XX
PI Fischer H, Illek B;
XX
DR WPI; 2002-105224/14.
DR P-PSDB; AAU74516.
XX
PT Pharmaceutical composition for the treatment of cystic fibrosis comprises
PT flavones or isoflavones.
XX
PS Disclosure; Col 31-38; 50pp; English.
XX
CC The invention relates to a pharmaceutical composition comprising one or
CC more compounds such as flavones or isoflavones, capable of stimulating
CC chloride transport in epithelial tissues, for treatment of cystic
CC fibrosis and other diseases associated with excessive accumulation of
CC mucus, e.g. chronic bronchitis and asthma. The active compound increases
CC expression of a cystic fibrosis transmembrane conductance regulator
CC (CFTR) in an epithelial cell and/or acts as a chemical chaperone that
CC increases trafficking of a CFTR to a plasma membrane in an epithelial
CC cell. This sequence represents DNA encoding the human delta-F508-CFTR
CC mutant polypeptide of the invention
XX
SQ Sequence 6126 BP; 1886 A; 1181 C; 1330 G; 1729 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 6; Length 6126;
Best Local Similarity 63.6%; Pred. No. 0.83;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| ||||| || | | | || | || | ||
Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 1604
Qy 65 TTGTCACTTCCGAGGA 81
|| ||| ||| | |||
Db 1605 TTCTCAGTTTTCTGGA 1621

RESULT 41

ADA37386

ID ADA37386 standard; DNA; 6126 BP.

XX

AC ADA37386;

XX

DT 20-NOV-2003 (first entry)

XX

DE DNA encoding human CFTR F508 deletion mutant.

XX

KW ds; gene; cystic fibrosis; chloride transport enhancement;

KW epithelial cell; airway epithelial cell; intestinal epithelial cell;

KW human; cystic fibrosis transmembrane conductance regulator; CFTR; mutant.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 133..4575

FT /*tag= a
 FT /product= "CFTR F508 deletion mutant"
 XX
 PN US2003096762-A1.
 XX
 PD 22-MAY-2003.
 XX
 PF 17-OCT-2001; 2001US-00982315.
 XX
 PR 16-OCT-1998; 98US-00174077.
 XX
 PA (CHIL-) CHILDRENS HOSPITAL OAKLAND.
 XX
 PI Fischer H, Illek B;
 XX
 DR WPI; 2003-616312/58.
 DR P-PSDB; ADA37387.
 XX
 PT Treating cystic fibrosis in a mammal, by administering flavones or
 PT isoflavones which stimulate chloride secretion, or by administering
 PT compounds such as resveratrol, ascorbic acid, ascorbate salts or
 PT dehydroascorbic acid.
 XX
 PS Disclosure; Page 17-20; 34pp; English.
 XX
 CC The invention relates to a method of treating cystic fibrosis in a
 CC mammal. The method is useful for treating cystic fibrosis in a mammal and
 CC for enhancing chloride transport in epithelial cells, preferably airway
 CC epithelial cells or intestinal epithelial cells present in mammals or
 CC epithelial cells present in pancreas, gallbladder, sweat duct, salivary
 CC gland or mammary epithelial cells. The present sequence represents the
 CC human cystic fibrosis transmembrane conductance regulator, CFTR, F508
 CC deletion mutant.
 XX
 SQ Sequence 6126 BP; 1886 A; 1181 C; 1330 G; 1729 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 8; Length 6126;
 Best Local Similarity 63.6%; Pred. No. 0.83;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || ||| |||| || | | | || | | || | ||
 Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 1604
 Qy 65 TTGTCACCTTCCGAGGA 81
 || ||| ||| | |||
 Db 1605 TTCTCAGTTTTCCTGGA 1621

RESULT 42

AAQ11371

ID AAQ11371 standard; DNA; 6127 BP.

XX

AC AAQ11371;

XX

DT 25-MAR-2003 (revised)

DT 22-MAY-1991 (first entry)

```

XX
DE   Mutant cystic fibrosis gene.
XX
KW   Cystis fibrosis; transmembrane conductance regulatory protein; CFTR;
KW   diagnosis; mutant; ss.
XX
OS   Homo sapiens.
XX
FH   Key          Location/Qualifiers
FT   exon         1. .185
FT               /*tag= c
FT               /number= 1
FT   misc_signal  12. .13
FT               /*tag= a
FT               /label= transcription_initiation_site
FT               /note= "by primer extension analysis"
FT   exon         186. .296
FT               /*tag= d
FT               /number= 2
FT   exon         297. .405
FT               /*tag= e
FT               /number= 3
FT   exon         406. .621
FT               /*tag= f
FT               /number= 4
FT   exon         622. .711
FT               /*tag= g
FT               /number= 5
FT   exon         712. .1001
FT               /*tag= h
FT               /number= 6
FT   exon         1002. .1248
FT               /*tag= i
FT               /number= 7
FT   exon         1249. .1341
FT               /*tag= j
FT               /number= 8
FT   exon         1342. .1523
FT               /*tag= k
FT               /number= 9
FT   exon         1524. .1712
FT               /*tag= l
FT               /number= 10
FT   exon         1713. .1807
FT               /*tag= m
FT               /number= 11
FT   exon         1809. .1894
FT               /*tag= n
FT               /number= 12
FT   exon         1895. .2617
FT               /*tag= o
FT               /number= 13
FT   exon         2618. .2785
FT               /*tag= p
FT               /number= 14
FT   exon         2786. .3036
FT               /*tag= q

```

FT /number= 15
 FT exon 3037. .3495
 FT /*tag= r
 FT /number= 16
 FT exon 3496. .3595
 FT /*tag= s
 FT /number= 17
 FT exon 3596. .3845
 FT /*tag= t
 FT /number= 18
 FT exon 3846. .4091
 FT /*tag= u
 FT /number= 19
 FT exon 4092. .4264
 FT /*tag= v
 FT /number= 20
 FT exon 4265. .4370
 FT /*tag= w
 FT /number= 21
 FT exon 4371. .6126
 FT /*tag= x
 FT /number= 22
 FT polyA_site 6126. .6126
 FT /*tag= b
 XX
 PN WO9102796-A.
 XX
 PD 07-MAR-1991.
 XX
 PF 22-AUG-1989; 89US-00396894.
 XX
 PR 22-AUG-1989; 89US-00396894.
 PR 24-AUG-1989; 89US-00399945.
 PR 31-AUG-1989; 89US-00401609.
 XX
 PA (HSCR-) HSC RES DEV CORP.
 PA (UNMI) UNIV MICHIGAN.
 XX
 PI Tsul LC, Riordan JR, Collins FS, Rommens JM, Jannuzzi MC;
 PI Kerem BS, Drumm ML, Buckwald M;
 XX
 DR WPI; 1991-087280/12.
 DR P-PSDB; AAR11602.
 XX
 PT Cystic fibrosis gene - used to produce prods. for screening, detection,
 PT diagnosis, therapy and studying cystic fibrosis.
 XX
 PS Disclosure; Fig 1;.163pp; English.
 XX
 CC The 3 bp CTT deletion at position 1653-1655 of the normal gene results in
 CC Phe 508 deletion in the amino acid sequence. The CF gene and its gene
 CC prod., nucleic acid probes and antibodies to the gene prod. can be used
 CC for screening and detection of CF carriers, CF diagnosis, prenatal CF
 CC screening and diagnosis, and gene and drug therapy. The prods. can also
 CC be used to develop improved methods of treatment and to study the
 CC disease. See AAQ11046 for the normal CF gene and AAQ11047-48 for CF
 CC probes. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-

CC 2003 to correct PI field.)

XX

SQ Sequence 6127 BP; 1887 A; 1181 C; 1329 G; 1730 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 2; Length 6127;

Best Local Similarity 63.6%; Pred. No. 0.83;

Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64

|| | |||| || ||| ||||| || | | | ||| || | ||| | |||

Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGGGAAGAATTTCATTCTG 1604

QY 65 TTGTCACTTTCCGAGGA 81

|| ||| ||| | |||

Db 1605 TTCTCAGTTTTCCTGGA 1621

RESULT 43

AAQ13068

ID AAQ13068 standard; DNA; 6128 BP.

XX

AC AAQ13068;

XX

DT 14-OCT-1991 (first entry)

XX

DE CFTR 556 del A.

XX

KW Deletion; mutant; diagnosis; antibodies; drug therapy; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 133. .4571

FT /*tag= a

FT /label= CFTR-mutant

FT misc_feature 185. .186

FT /*tag= b

FT /label= exon_junction

FT misc_feature 296. .297

FT /*tag= c

FT /label= exon_junction

FT misc_feature 372. .438

FT /*tag= w

FT /label= membrane-spanning_segment

FT misc_feature 405. .406

FT /*tag= d

FT /label= exon_junction

FT misc_feature 484. .545

FT /*tag= x

FT /label= membrane-spanning_segment

FT misc_feature 620. .621

FT /*tag= e

FT /label= exon_junction

FT misc_feature 710. .711

FT /*tag= f

FT /label= exon_junction

FT misc_feature 713. .776

```

FT          /*tag= y
FT          /label= membrane-spanning_segment
FT  misc_feature  792. .854
FT          /*tag= z
FT          /label= membrane-spanning_segment
FT  misc_feature  1000. .1001
FT          /*tag= g
FT          /label= exon_junction
FT  misc_feature  1053. .1115
FT          /*tag= aa
FT          /label= membrane-spanning_segment
FT  misc_feature  1119. .1181
FT          /*tag= ab
FT          /label= membrane-spanning_segment
FT  misc_feature  1247. .1248
FT          /*tag= h
FT          /label= exon_junction
FT  misc_feature  1340. .1341
FT          /*tag= i
FT          /label= exon_junction
FT  misc_binding  1428. .1883
FT          /*tag= ai
FT          /label= ATP-binding_fold
FT  misc_feature  1522. .1523
FT          /*tag= j
FT          /label= exon_junction
FT  misc_feature  1715. .1716
FT          /*tag= k
FT          /label= exon_junction
FT  misc_feature  1810. .1811
FT          /*tag= l
FT          /label= exon_junction
FT  misc_feature  1897. .1898
FT          /*tag= m
FT          /label= exon_junction
FT  misc_feature  2621. .2622
FT          /*tag= n
FT          /label= exon_junction
FT  misc_feature  2709. .2771
FT          /*tag= ac
FT          /label= membrane-spanning_segment
FT  misc_feature  2788. .2789
FT          /*tag= o
FT          /label= exon_junction
FT  misc_feature  2865. .2927
FT          /*tag= ad
FT          /label= membrane-spanning_segment
FT  misc_feature  3039. .3040
FT          /*tag= p
FT          /label= exon_junction
FT  misc_feature  3102. .3164
FT          /*tag= ae
FT          /label= membrane-spanning_segment
FT  misc_feature  3170. .3233
FT          /*tag= af
FT          /label= membrane-spanning_segment
FT  misc_feature  3438. .3500

```

```

FT          /*tag= ag
FT          /label= membrane-spanning_segment
FT  misc_feature  3498. .3499
FT          /*tag= q
FT          /label= exon_junction
FT  misc_feature  3516. .3581
FT          /*tag= ah
FT          /label= membrane-spanning_segment
FT  misc_feature  3598. .3599
FT          /*tag= r
FT          /label= exon_junction
FT  misc_binding  3786. .4289
FT          /*tag= aj
FT          /label= ATP-binding_fold
FT  misc_feature  3848. .3849
FT          /*tag= s
FT          /label= exon_junction
FT  misc_feature  4094. .4095
FT          /*tag= t
FT          /label= exon_junction
FT  misc_feature  4267. .4268
FT          /*tag= u
FT          /label= exon_junction
FT  misc_feature  4373. .4374
FT          /*tag= v
FT          /label= exon_junction
XX
PN  WO9110734-A.
XX
PD  25-JUL-1991.
XX
PF  12-JAN-1990;  90CA-02007699.
XX
PR  12-JAN-1990;  90CA-02007699.
PR  01-MAR-1990;  90CA-02011253.
PR  10-JUL-1990;  90CA-02020817.
XX
PA  (HSCR-) HSC RES DEV CORP.
XX
PI  Tsui LC,  Rommens JM,  Kerem B;
XX
DR  WPI; 1991-238022/32.
DR  P-PSDB; AAR13304.
XX
PT  Mutant cystic fibrosis trans-membrane conductance regulator gene - used
PT  for producing prods. for diagnosis, screening and therapy of cystic
PT  fibrosis.
XX
PS  Claim 2; Page 121; 178pp; English.
XX
CC  556 del A is a frameshift mutation in exon 4. The mutant CF gene when
CC  expressed in cells of the human body, is associated with altered cell
CC  function which correlates with the genetic disease cystic fibrosis. See
CC  also AAQ13053-72
XX
SQ  Sequence 6128 BP; 1884 A; 1183 C; 1329 G; 1732 T; 0 U; 0 Other;

```

Query Match 31.6%; Score 32.2; DB 2; Length 6128;
 Best Local Similarity 63.6%; Pred. No. 0.83;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
        || | |||| ||   ||| ||||| ||   | | | ||| || | ||| | || |||
Db      1544 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTCATTCTG 1603

Qy      65 TTGTCACTTTCCGAGGA 81
        || ||| ||| | |||
Db      1604 TTCTCAGTTTTCCTGGA 1620
  
```

RESULT 44

AAQ13072

ID AAQ13072 standard; DNA; 6128 BP.

XX

AC AAQ13072;

XX

DT 14-OCT-1991 (first entry)

XX

DE CFTR 3659 del C.

XX

KW Deletion; mutant; diagnosis; antibodies; drug therapy; ss.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
FT	CDS	133. .4572
FT		/*tag= a
FT		/label= CFTR-mutant
FT	misc_feature	185. .186
FT		/*tag= b
FT		/label= exon_junction
FT	misc_feature	296. .297
FT		/*tag= c
FT		/label= exon_junction
FT	misc_feature	372. .438
FT		/*tag= w
FT		/label= membrane-spanning_segment
FT	misc_feature	405. .406
FT		/*tag= d
FT		/label= exon_junction
FT	misc_feature	484. .546
FT		/*tag= x
FT		/label= membrane-spanning_segment
FT	misc_feature	621. .622
FT		/*tag= e
FT		/label= exon_junction
FT	misc_feature	711. .712
FT		/*tag= f
FT		/label= exon_junction
FT	misc_feature	714. .777
FT		/*tag= y
FT		/label= membrane-spanning_segment
FT	misc_feature	793. .855
FT		/*tag= z

FT		/label= membrane-spanning_segment
FT	misc_feature	1001. .1002
FT		/*tag= g
FT		/label= exon_junction
FT	misc_feature	1054. .1116
FT		/*tag= aa
FT		/label= membrane-spanning_segment
FT	misc_feature	1120. .1182
FT		/*tag= ab
FT		/label= membrane-spanning_segment
FT	misc_feature	1248. .1249
FT		/*tag= h
FT		/label= exon_junction
FT	misc_feature	1341. .1342
FT		/*tag= i
FT		/label= exon_junction
FT	misc_binding	1429. .1884
FT		/*tag= ai
FT		/label= ATP-binding_fold
FT	misc_feature	1523. .1524
FT		/*tag= j
FT		/label= exon_junction
FT	misc_feature	1716. .1717
FT		/*tag= k
FT		/label= exon_junction
FT	misc_feature	1811. .1812
FT		/*tag= l
FT		/label= exon_junction
FT	misc_feature	1898. .1899
FT		/*tag= m
FT		/label= exon_junction
FT	misc_feature	2622. .2623
FT		/*tag= n
FT		/label= exon_junction
FT	misc_feature	2710. .2772
FT		/*tag= ac
FT		/label= membrane-spanning_segment
FT	misc_feature	2789. .2790
FT		/*tag= o
FT		/label= exon_junction
FT	misc_feature	2866. .2928
FT		/*tag= ad
FT		/label= membrane-spanning_segment
FT	misc_feature	3040. .3041
FT		/*tag= p
FT		/label= exon_junction
FT	misc_feature	3103. .3165
FT		/*tag= ae
FT		/label= membrane-spanning_segment
FT	misc_feature	3171. .3234
FT		/*tag= af
FT		/label= membrane-spanning_segment
FT	misc_feature	3439. .3501
FT		/*tag= ag
FT		/label= membrane-spanning_segment
FT	misc_feature	3499. .3500
FT		/*tag= q

FT /label= exon_junction
 FT misc_feature 3517. .3582
 FT /*tag= ah
 FT /label= membrane-spanning_segment
 FT misc_feature 3599. .3600
 FT /*tag= r
 FT /label= exon_junction
 FT misc_binding 3786. .4289
 FT /*tag= aj
 FT /label= ATP-binding_fold
 FT misc_feature 3848. .3849
 FT /*tag= s
 FT /label= exon_junction
 FT misc_feature 4094. .4095
 FT /*tag= t
 FT /label= exon_junction
 FT misc_feature 4267. .4268
 FT /*tag= u
 FT /label= exon_junction
 FT misc_feature 4373. .4374
 FT /*tag= v
 FT /label= exon_junction

XX

PN WO9110734-A.

XX

PD 25-JUL-1991.

XX

PF 12-JAN-1990; 90CA-02007699.

XX

PR 12-JAN-1990; 90CA-02007699.

PR 01-MAR-1990; 90CA-02011253.

PR 10-JUL-1990; 90CA-02020817.

XX

PA (HSCR-) HSC RES DEV CORP.

XX

PI Tsui LC, Rommens JM, Kerem B;

XX

DR WPI; 1991-238022/32.

DR P-PSDB; AAR13308.

XX

PT Mutant cystic fibrosis trans-membrane conductance regulator gene - used
 PT for producing prods. for diagnosis, screening and therapy of cystic
 PT fibrosis.

XX

PS Claim 2; Page 121; 178pp; English.

XX

CC 3659 del C is a frameshift mutation in exon 19. The 3659 del C mutation
 CC results in a shortened polypeptide significantly different from the
 CC single amino acid deletions or alterations. The mutant CF gene when
 CC expressed in cells of the human body, is associated with altered cell
 CC function which correlates with the genetic disease cystic fibrosis. See
 CC also AAQ13053-72

XX

SQ Sequence 6128 BP; 1885 A; 1182 C; 1329 G; 1732 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 2; Length 6128;
 Best Local Similarity 63.6%; Pred. No. 0.83;

Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
QY          5 TAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
              || | ||| ||   ||| |||| ||   || | ||| || |   ||| | ||   |||
Db          1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTATTCTG 1604

QY          65 TTGTCACCTTTCCGAGGA 81
              || ||| ||| |   |||
Db          1605 TTCTCAGTTTTTCCTGGA 1621
```

RESULT 45

AAQ13056

ID AAQ13056 standard; DNA; 6129 BP.

XX

AC AAQ13056;

XX

DT 14-OCT-1991 (first entry)

XX

DE CFTR G178R.

XX

KW Deletion; mutant; diagnosis; antibodies; drug therapy; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 133. .4572

FT /*tag= a

FT /label= CFTR-mutant

FT misc_feature 185. .186

FT /*tag= b

FT /label= exon_junction

FT misc_feature 296. .297

FT /*tag= c

FT /label= exon_junction

FT misc_feature 372. .438

FT /*tag= w

FT /label= membrane-spanning_segment

FT misc_feature 405. .406

FT /*tag= d

FT /label= exon_junction

FT misc_feature 484. .546

FT /*tag= x

FT /label= membrane-spanning_segment

FT misc_feature 621. .622

FT /*tag= e

FT /label= exon_junction

FT misc_feature 711. .712

FT /*tag= f

FT /label= exon_junction

FT misc_feature 714. .777

FT /*tag= y

FT /label= membrane-spanning_segment

FT misc_feature 793. .855

FT /*tag= z

FT /label= membrane-spanning_segment

FT misc_feature 1001. .1002

```

FT          /*tag= g
FT          /label= exon_junction
FT  misc_feature  1054. .1116
FT          /*tag= aa
FT          /label= membrane-spanning_segment
FT  misc_feature  1120. .1182
FT          /*tag= ab
FT          /label= membrane-spanning_segment
FT  misc_feature  1248. .1249
FT          /*tag= h
FT          /label= exon_junction
FT  misc_feature  1341. .1342
FT          /*tag= i
FT          /label= exon_junction
FT  misc_binding  1429. .1884
FT          /*tag= ai
FT          /label= ATP-binding_fold
FT  misc_feature  1523. .1524
FT          /*tag= j
FT          /label= exon_junction
FT  misc_feature  1716. .1717
FT          /*tag= k
FT          /label= exon_junction
FT  misc_feature  1811. .1812
FT          /*tag= l
FT          /label= exon_junction
FT  misc_feature  1898. .1899
FT          /*tag= m
FT          /label= exon_junction
FT  misc_feature  2622. .2623
FT          /*tag= n
FT          /label= exon_junction
FT  misc_feature  2710. .2772
FT          /*tag= ac
FT          /label= membrane-spanning_segment
FT  misc_feature  2789. .2790
FT          /*tag= o
FT          /label= exon_junction
FT  misc_feature  2866. .2928
FT          /*tag= ad
FT          /label= membrane-spanning_segment
FT  misc_feature  3040. .3041
FT          /*tag= p
FT          /label= exon_junction
FT  misc_feature  3103. .3165
FT          /*tag= ae
FT          /label= membrane-spanning_segment
FT  misc_feature  3171. .3234
FT          /*tag= af
FT          /label= membrane-spanning_segment
FT  misc_feature  3439. .3501
FT          /*tag= ag
FT          /label= membrane-spanning_segment
FT  misc_feature  3499. .3500
FT          /*tag= q
FT          /label= exon_junction
FT  misc_feature  3517. .3582

```

```

FT          /*tag= ah
FT          /label= membrane-spanning_segment
FT  misc_feature  3599. .3600
FT          /*tag= r
FT          /label= exon_junction
FT  misc_binding  3787. .4290
FT          /*tag= aj
FT          /label= ATP-binding_fold
FT  misc_feature  3849. .3850
FT          /*tag= s
FT          /label= exon_junction
FT  misc_feature  4095. .4096
FT          /*tag= t
FT          /label= exon_junction
FT  misc_feature  4268. .4269
FT          /*tag= u
FT          /label= exon_junction
FT  misc_feature  4374. .4375
FT          /*tag= v
FT          /label= exon_junction

```

```

XX
PN  WO9110734-A.
XX
PD  25-JUL-1991.
XX
PF  12-JAN-1990;  90CA-02007699.
XX
PR  12-JAN-1990;  90CA-02007699.
PR  01-MAR-1990;  90CA-02011253.
PR  10-JUL-1990;  90CA-02020817.

```

```

XX
PA  (HSCR-) HSC RES DEV CORP.
XX
PI  Tsui LC,  Rommens JM,  Kerem B;
XX
DR  WPI; 1991-238022/32.
DR  P-PSDB; AAR13234.

```

```

XX
PT  Mutant cystic fibrosis trans-membrane conductance regulator gene - used
PT  for producing prods. for diagnosis, screening and therapy of cystic
PT  fibrosis.

```

```

XX
PS  Claim 1; Page 121; 178pp; English.
XX

```

```

CC  The G178R mutation in exon 5 involves a G to A transition at nucleotide
CC  position 664. Nucleotide 6129 is followed by a poly (dA) tract. The
CC  mutant CF gene when expressed in cells of the human body, is associated
CC  with altered cell function which correlates with the genetic disease
CC  cystic fibrosis. See also AAQ13053-72

```

```

XX
SQ  Sequence 6129 BP; 1886 A; 1183 C; 1328 G; 1732 T; 0 U; 0 Other;

```

```

Query Match          31.6%;  Score 32.2;  DB 2;  Length 6129;
Best Local Similarity 63.6%;  Pred. No. 0.83;
Matches  49;  Conservative  0;  Mismatches  28;  Indels  0;  Gaps  0;

```

```

QY          5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64

```

```

      || | |||| ||   ||| ||||| ||   | | | ||| || |   ||| | ||   |||
Db      1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTCATTCTG 1604

Qy      65 TTGTCACCTTTCCGAGGA 81
      || ||| ||| |   |||
Db      1605 TTCTCAGTTTTCTCTGGA 1621

```

RESULT 46

AAQ13060

ID AAQ13060 standard; DNA; 6129 BP.

XX

AC AAQ13060;

XX

DT 14-OCT-1991 (first entry)

XX

DE CFTR S549R.

XX

KW Deletion; mutant; diagnosis; antibodies; drug therapy; ss.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
FT	CDS	133. .4572
FT		/*tag= a
FT		/label= CFTR-mutant
FT	misc_feature	185. .186
FT		/*tag= b
FT		/label= exon_junction
FT	misc_feature	296. .297
FT		/*tag= c
FT		/label= exon_junction
FT	misc_feature	372. .438
FT		/*tag= w
FT		/label= membrane-spanning_segment
FT	misc_feature	405. .406
FT		/*tag= d
FT		/label= exon_junction
FT	misc_feature	484. .546
FT		/*tag= x
FT		/label= membrane-spanning_segment
FT	misc_feature	621. .622
FT		/*tag= e
FT		/label= exon_junction
FT	misc_feature	711. .712
FT		/*tag= f
FT		/label= exon_junction
FT	misc_feature	714. .777
FT		/*tag= y
FT		/label= membrane-spanning_segment
FT	misc_feature	793. .855
FT		/*tag= z
FT		/label= membrane-spanning_segment
FT	misc_feature	1001. .1002
FT		/*tag= g
FT		/label= exon_junction
FT	misc_feature	1054. .1116

```

FT          /*tag= aa
FT          /label= membrane-spanning_segment
FT  misc_feature  1120. .1182
FT          /*tag= ab
FT          /label= membrane-spanning_segment
FT  misc_feature  1248. .1249
FT          /*tag= h
FT          /label= exon_junction
FT  misc_feature  1341. .1342
FT          /*tag= i
FT          /label= exon_junction
FT  misc_binding  1429. .1884
FT          /*tag= ai
FT          /label= ATP-binding_fold
FT  misc_feature  1523. .1524
FT          /*tag= j
FT          /label= exon_junction
FT  misc_feature  1716. .1717
FT          /*tag= k
FT          /label= exon_junction
FT  misc_feature  1811. .1812
FT          /*tag= l
FT          /label= exon_junction
FT  misc_feature  1898. .1899
FT          /*tag= m
FT          /label= exon_junction
FT  misc_feature  2622. .2623
FT          /*tag= n
FT          /label= exon_junction
FT  misc_feature  2710. .2772
FT          /*tag= ac
FT          /label= membrane-spanning_segment
FT  misc_feature  2789. .2790
FT          /*tag= o
FT          /label= exon_junction
FT  misc_feature  2866. .2928
FT          /*tag= ad
FT          /label= membrane-spanning_segment
FT  misc_feature  3040. .3041
FT          /*tag= p
FT          /label= exon_junction
FT  misc_feature  3103. .3165
FT          /*tag= ae
FT          /label= membrane-spanning_segment
FT  misc_feature  3171. .3234
FT          /*tag= af
FT          /label= membrane-spanning_segment
FT  misc_feature  3439. .3501
FT          /*tag= ag
FT          /label= membrane-spanning_segment
FT  misc_feature  3499. .3500
FT          /*tag= q
FT          /label= exon_junction
FT  misc_feature  3517. .3582
FT          /*tag= ah
FT          /label= membrane-spanning_segment
FT  misc_feature  3599. .3600

```

```

FT          /*tag= r
FT          /label= exon_junction
FT  misc_binding  3787. .4290
FT          /*tag= aj
FT          /label= ATP-binding_fold
FT  misc_feature  3849. .3850
FT          /*tag= s
FT          /label= exon_junction
FT  misc_feature  4095. .4096
FT          /*tag= t
FT          /label= exon_junction
FT  misc_feature  4268. .4269
FT          /*tag= u
FT          /label= exon_junction
FT  misc_feature  4374. .4375
FT          /*tag= v
FT          /label= exon_junction
XX
PN  WO9110734-A.
XX
PD  25-JUL-1991.
XX
PF  12-JAN-1990;  90CA-02007699.
XX
PR  12-JAN-1990;  90CA-02007699.
PR  01-MAR-1990;  90CA-02011253.
PR  10-JUL-1990;  90CA-02020817.
XX
PA  (HSCR-) HSC RES DEV CORP.
XX
PI  Tsui LC,  Rommens JM,  Kerem B;
XX
DR  WPI; 1991-238022/32.
DR  P-PSDB; AAR13297.
XX
PT  Mutant cystic fibrosis trans-membrane conductance regulator gene - used
PT  for producing prods. for diagnosis, screening and therapy of cystic
PT  fibrosis.
XX
PS  Claim 1; Page 121; 178pp; English.
XX
CC  In the S549R, the highly conserved Ser of the nucleotide binding domain
CC  at position 549 is changed to Arg. The codon change is AGT to AGG. The
CC  mutant CF gene when expressed in cells of the human body, is associated
CC  with altered cell function which correlates with the genetic disease
CC  cystic fibrosis. See also AAQ13053-72
XX
SQ  Sequence 6129 BP; 1885 A; 1183 C; 1330 G; 1731 T; 0 U; 0 Other;

Query Match          31.6%;  Score 32.2;  DB 2;  Length 6129;
Best Local Similarity 63.6%;  Pred. No. 0.83;
Matches  49;  Conservative  0;  Mismatches  28;  Indels  0;  Gaps  0;

Qy          5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
           || | ||| ||   ||| |||| ||   | | | || | || | ||| || | |||
Db        1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTCATTCTG 1604

```


Qy 65 TTGTCACCTTTCCGAGGA 81
 || ||| ||| | |||
 Db 1605 TTCTCAGTTTTCTGGA 1621

RESULT 47

AAQ13071

ID AAQ13071 standard; DNA; 6129 BP.

XX

AC AAQ13071;

XX

DT 14-OCT-1991 (first entry)

XX

DE CFTR 1717 -1G -> A.

XX

KW Deletion; mutant; diagnosis; antibodies; drug therapy; ss.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
FT	CDS	133. .4572
FT		/*tag= a
FT		/label= CFTR-mutant
FT	misc_feature	185. .186
FT		/*tag= b
FT		/label= exon_junction
FT	misc_feature	296. .297
FT		/*tag= c
FT		/label= exon_junction
FT	misc_feature	372. .438
FT		/*tag= w
FT		/label= membrane-spanning_segment
FT	misc_feature	405. .406
FT		/*tag= d
FT		/label= exon_junction
FT	misc_feature	484. .546
FT		/*tag= x
FT		/label= membrane-spanning_segment
FT	misc_feature	621. .622
FT		/*tag= e
FT		/label= exon_junction
FT	misc_feature	711. .712
FT		/*tag= f
FT		/label= exon_junction
FT	misc_feature	714. .777
FT		/*tag= y
FT		/label= membrane-spanning_segment
FT	misc_feature	793. .855
FT		/*tag= z
FT		/label= membrane-spanning_segment
FT	misc_feature	1001. .1002
FT		/*tag= g
FT		/label= exon_junction
FT	misc_feature	1054. .1116
FT		/*tag= aa
FT		/label= membrane-spanning_segment
FT	misc_feature	1120. .1182

```

FT          /*tag= ab
FT          /label= membrane-spanning_segment
FT  misc_feature  1248. .1249
FT          /*tag= h
FT          /label= exon_junction
FT  misc_feature  1341. .1342
FT          /*tag= i
FT          /label= exon_junction
FT  misc_binding  1429. .1884
FT          /*tag= ai
FT          /label= ATP-binding_fold
FT  misc_feature  1523. .1524
FT          /*tag= j
FT          /label= exon_junction
FT  misc_feature  1716. .1717
FT          /*tag= k
FT          /label= exon_junction
FT  misc_feature  1811. .1812
FT          /*tag= l
FT          /label= exon_junction
FT  misc_feature  1898. .1899
FT          /*tag= m
FT          /label= exon_junction
FT  misc_feature  2622. .2623
FT          /*tag= n
FT          /label= exon_junction
FT  misc_feature  2710. .2772
FT          /*tag= ac
FT          /label= membrane-spanning_segment
FT  misc_feature  2789. .2790
FT          /*tag= o
FT          /label= exon_junction
FT  misc_feature  2866. .2928
FT          /*tag= ad
FT          /label= membrane-spanning_segment
FT  misc_feature  3040. .3041
FT          /*tag= p
FT          /label= exon_junction
FT  misc_feature  3103. .3165
FT          /*tag= ae
FT          /label= membrane-spanning_segment
FT  misc_feature  3171. .3234
FT          /*tag= af
FT          /label= membrane-spanning_segment
FT  misc_feature  3439. .3501
FT          /*tag= ag
FT          /label= membrane-spanning_segment
FT  misc_feature  3499. .3500
FT          /*tag= q
FT          /label= exon_junction
FT  misc_feature  3517. .3582
FT          /*tag= ah
FT          /label= membrane-spanning_segment
FT  misc_feature  3599. .3600
FT          /*tag= r
FT          /label= exon_junction
FT  misc_binding  3787. .4290

```

FT /*tag= aj
 FT /label= ATP-binding_fold
 FT misc_feature 3849. .3850
 FT /*tag= s
 FT /label= exon_junction
 FT misc_feature 4095. .4096
 FT /*tag= t
 FT /label= exon_junction
 FT misc_feature 4268. .4269
 FT /*tag= u
 FT /label= exon_junction
 FT misc_feature 4374. .4375
 FT /*tag= v
 FT /label= exon_junction

XX

PN WO9110734-A.

XX

PD 25-JUL-1991.

XX

PF 12-JAN-1990; 90CA-02007699.

XX

PR 12-JAN-1990; 90CA-02007699.

PR 01-MAR-1990; 90CA-02011253.

PR 10-JUL-1990; 90CA-02020817.

XX

PA (HSCR-) HSC RES DEV CORP.

XX

PI Tsui LC, Rommens JM, Kerem B;

XX

DR WPI; 1991-238022/32.

XX

PT Mutant cystic fibrosis trans-membrane conductance regulator gene - used
 PT for producing prods. for diagnosis, screening and therapy of cystic
 PT fibrosis.

XX

PS Claim 2; Page 121; 178pp; English.

XX

CC In the 1717 -1G -> A mutation a putative plice mutation is found in front
 CC of exon 11. This mutation is located at the last nucleotide of the intron
 CC before exon 11, and is predicted to lead to polypeptides which cannot be
 CC as yet exactly defined. Nucleotide 6129 is followed by a poly (dA) tract.
 CC The mutant CF gene when expressed in cells of the human body, is
 CC associated with altered cell function which correlates with the genetic
 CC disease cystic fibrosis. See also AAQ13053-72

XX

SQ Sequence 6129 BP; 1886 A; 1183 C; 1328 G; 1732 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 2; Length 6129;
 Best Local Similarity 63.6%; Pred. No. 0.83;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 ||| ||||| || ||| ||||| || | | | ||| || | ||| ||| |||
 Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGGGAAGAATTCATTCTG 1604
 Qy 65 TTGTCACTTCCGAGGA 81
 || ||| ||| | |||

Db 1605 TTCTCAGTTTTCCTGGA 1621

RESULT 48

AAQ13054

ID AAQ13054 standard; DNA; 6129 BP.

XX

AC AAQ13054;

XX

DT 14-OCT-1991 (first entry)

XX

DE CFTR G85E.

XX

KW Deletion; mutant; diagnosis; antibodies; drug therapy; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 133. .4572

FT /*tag= a

FT /label= CFTR-mutant

FT misc_feature 185. .186

FT /*tag= b

FT /label= exon_junction

FT misc_feature 296. .297

FT /*tag= c

FT /label= exon_junction

FT misc_feature 372. .438

FT /*tag= w

FT /label= membrane-spanning_segment

FT misc_feature 405. .406

FT /*tag= d

FT /label= exon_junction

FT misc_feature 484. .546

FT /*tag= x

FT /label= membrane-spanning_segment

FT misc_feature 621. .622

FT /*tag= e

FT /label= exon_junction

FT misc_feature 711. .712

FT /*tag= f

FT /label= exon_junction

FT misc_feature 714. .777

FT /*tag= y

FT /label= membrane-spanning_segment

FT misc_feature 793. .855

FT /*tag= z

FT /label= membrane-spanning_segment

FT misc_feature 1001. .1002

FT /*tag= g

FT /label= exon_junction

FT misc_feature 1054. .1116

FT /*tag= aa

FT /label= membrane-spanning_segment

FT misc_feature 1120. .1182

FT /*tag= ab

FT /label= membrane-spanning_segment

```

FT      misc_feature      1248. .1249
FT      /*tag= h
FT      /label= exon_junction
FT      misc_feature      1341. .1342
FT      /*tag= i
FT      /label= exon_junction
FT      misc_binding      1429. .1884
FT      /*tag= ai
FT      /label= ATP-binding_fold
FT      misc_feature      1523. .1524
FT      /*tag= j
FT      /label= exon_junction
FT      misc_feature      1716. .1717
FT      /*tag= k
FT      /label= exon_junction
FT      misc_feature      1811. .1812
FT      /*tag= l
FT      /label= exon_junction
FT      misc_feature      1898. .1899
FT      /*tag= m
FT      /label= exon_junction
FT      misc_feature      2622. .2623
FT      /*tag= n
FT      /label= exon_junction
FT      misc_feature      2710. .2772
FT      /*tag= ac
FT      /label= membrane-spanning_segment
FT      misc_feature      2789. .2790
FT      /*tag= o
FT      /label= exon_junction
FT      misc_feature      2866. .2928
FT      /*tag= ad
FT      /label= membrane-spanning_segment
FT      misc_feature      3040. .3041
FT      /*tag= p
FT      /label= exon_junction
FT      misc_feature      3103. .3165
FT      /*tag= ae
FT      /label= membrane-spanning_segment
FT      misc_feature      3171. .3234
FT      /*tag= af
FT      /label= membrane-spanning_segment
FT      misc_feature      3439. .3501
FT      /*tag= ag
FT      /label= membrane-spanning_segment
FT      misc_feature      3499. .3500
FT      /*tag= q
FT      /label= exon_junction
FT      misc_feature      3517. .3582
FT      /*tag= ah
FT      /label= membrane-spanning_segment
FT      misc_feature      3599. .3600
FT      /*tag= r
FT      /label= exon_junction
FT      misc_binding      3787. .4290
FT      /*tag= aj
FT      /label= ATP-binding_fold

```

FT misc_feature 3849. .3850
 FT /*tag= s
 FT /label= exon_junction
 FT misc_feature 4095. .4096
 FT /*tag= t
 FT /label= exon_junction
 FT misc_feature 4268. .4269
 FT /*tag= u
 FT /label= exon_junction
 FT misc_feature 4374. .4375
 FT /*tag= v
 FT /label= exon_junction

XX

PN WO9110734-A.

XX

PD 25-JUL-1991.

XX

PF 12-JAN-1990; 90CA-02007699.

XX

PR 12-JAN-1990; 90CA-02007699.

PR 01-MAR-1990; 90CA-02011253.

PR 10-JUL-1990; 90CA-02020817.

XX

PA (HSCR-) HSC RES DEV CORP.

XX

PI Tsui LC, Rommens JM, Kerem B;

XX

DR WPI; 1991-238022/32.

DR P-PSDB; AAR13232.

XX

PT Mutant cystic fibrosis trans-membrane conductance regulator gene - used
 PT for producing prods. for diagnosis, screening and therapy of cystic
 PT fibrosis.

XX

PS Claim 1; Page 121; 178pp; English.

XX

CC The G85E mutation in exon3 involves a G to A transition at nucleotide
 CC position 386. The predicted Gly to Glu amino acid change is associated
 CC with a group IIb haplotype. The mutation destroys a HinfI site.
 CC Nucleotide 6129 is followed by a poly (dA) tract. The mutant CF gene when
 CC expressed in cells of the human body, is associated with altered cell
 CC function which correlates with the genetic disease cystic fibrosis. See
 CC also AAQ13053-72

XX

SQ Sequence 6129 BP; 1886 A; 1183 C; 1328 G; 1732 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 2; Length 6129;

Best Local Similarity 63.6%; Pred. No. 0.83;

Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 ||| ||||| || ||| ||||| || | | | ||| || | ||| ||| |||

Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 1604

Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||

Db 1605 TTCTCAGTTTTCCTGGA 1621

RESULT 49

AAQ13065

ID AAQ13065 standard; DNA; 6129 BP.

XX

AC AAQ13065;

XX

DT 14-OCT-1991 (first entry)

XX

DE CFTR L1077P.

XX

KW Deletion; mutant; diagnosis; antibodies; drug therapy; ss.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
FT	CDS	133. .4572
FT		/*tag= a
FT		/label= CFTR-mutant
FT	misc_feature	185. .186
FT		/*tag= b
FT		/label= exon_junction
FT	misc_feature	296. .297
FT		/*tag= c
FT		/label= exon_junction
FT	misc_feature	372. .438
FT		/*tag= w
FT		/label= membrane-spanning_segment
FT	misc_feature	405. .406
FT		/*tag= d
FT		/label= exon_junction
FT	misc_feature	484. .546
FT		/*tag= x
FT		/label= membrane-spanning_segment
FT	misc_feature	621. .622
FT		/*tag= e
FT		/label= exon_junction
FT	misc_feature	711. .712
FT		/*tag= f
FT		/label= exon_junction
FT	misc_feature	714. .777
FT		/*tag= y
FT		/label= membrane-spanning_segment
FT	misc_feature	793. .855
FT		/*tag= z
FT		/label= membrane-spanning_segment
FT	misc_feature	1001. .1002
FT		/*tag= g
FT		/label= exon_junction
FT	misc_feature	1054. .1116
FT		/*tag= aa
FT		/label= membrane-spanning_segment
FT	misc_feature	1120. .1182
FT		/*tag= ab
FT		/label= membrane-spanning_segment
FT	misc_feature	1248. .1249

```

FT          /*tag= h
FT          /label= exon_junction
FT  misc_feature  1341. .1342
FT          /*tag= i
FT          /label= exon_junction
FT  misc_binding  1429. .1884
FT          /*tag= ai
FT          /label= ATP-binding_fold
FT  misc_feature  1523. .1524
FT          /*tag= j
FT          /label= exon_junction
FT  misc_feature  1716. .1717
FT          /*tag= k
FT          /label= exon_junction
FT  misc_feature  1811. .1812
FT          /*tag= l
FT          /label= exon_junction
FT  misc_feature  1898. .1899
FT          /*tag= m
FT          /label= exon_junction
FT  misc_feature  2622. .2623
FT          /*tag= n
FT          /label= exon_junction
FT  misc_feature  2710. .2772
FT          /*tag= ac
FT          /label= membrane-spanning_segment
FT  misc_feature  2789. .2790
FT          /*tag= o
FT          /label= exon_junction
FT  misc_feature  2866. .2928
FT          /*tag= ad
FT          /label= membrane-spanning_segment
FT  misc_feature  3040. .3041
FT          /*tag= p
FT          /label= exon_junction
FT  misc_feature  3103. .3165
FT          /*tag= ae
FT          /label= membrane-spanning_segment
FT  misc_feature  3171. .3234
FT          /*tag= af
FT          /label= membrane-spanning_segment
FT  misc_feature  3439. .3501
FT          /*tag= ag
FT          /label= membrane-spanning_segment
FT  misc_feature  3499. .3500
FT          /*tag= q
FT          /label= exon_junction
FT  misc_feature  3517. .3582
FT          /*tag= ah
FT          /label= membrane-spanning_segment
FT  misc_feature  3599. .3600
FT          /*tag= r
FT          /label= exon_junction
FT  misc_binding  3787. .4290
FT          /*tag= aj
FT          /label= ATP-binding_fold
FT  misc_feature  3849. .3850

```


FT /*tag= s
 FT /label= exon_junction
 FT misc_feature 4095. .4096
 FT /*tag= t
 FT /label= exon_junction
 FT misc_feature 4268. .4269
 FT /*tag= u
 FT /label= exon_junction
 FT misc_feature 4374. .4375
 FT /*tag= v
 FT /label= exon_junction

XX

PN WO9110734-A.

XX

PD 25-JUL-1991.

XX

PF 12-JAN-1990; 90CA-02007699.

XX

PR 12-JAN-1990; 90CA-02007699.

PR 01-MAR-1990; 90CA-02011253.

PR 10-JUL-1990; 90CA-02020817.

XX

PA (HSCR-) HSC RES DEV CORP.

XX

PI Tsui LC, Rommens JM, Kerem B;

XX

DR WPI; 1991-238022/32.

DR P-PSDB; AAR13302.

XX

PT Mutant cystic fibrosis trans-membrane conductance regulator gene - used
 PT for producing prods. for diagnosis, screening and therapy of cystic
 PT fibrosis.

XX

PS Claim 1; Page 121; 178pp; English.

XX

CC In the L1077P mutation a T to C change is detected at nucleotide position
 CC 3362. The mutant CF gene when expressed in cells of the human body, is
 CC associated with altered cell function which correlates with the genetic
 CC disease cystic fibrosis. See also AAQ13053-72

XX

SQ Sequence 6129 BP; 1885 A; 1184 C; 1329 G; 1731 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 2; Length 6129;
 Best Local Similarity 63.6%; Pred. No. 0.83;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || ||| ||||| || | | | || | || | |||
 Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG 1604
 Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 1605 TTCTCAGTTTTCCTGGA 1621

RESULT 50
 AAQ13063

ID AAQ13063 standard; DNA; 6129 BP.
 XX
 AC AAQ13063;
 XX
 DT 14-OCT-1991 (first entry)
 XX
 DE CFTR Y563N.
 XX
 KW Deletion; mutant; diagnosis; antibodies; drug therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 133. .4572
 FT /*tag= a
 FT /label= CFTR-mutant
 FT misc_feature 185. .186
 FT /*tag= b
 FT /label= exon_junction
 FT misc_feature 296. .297
 FT /*tag= c
 FT /label= exon_junction
 FT misc_feature 372. .438
 FT /*tag= w
 FT /label= membrane-spanning_segment
 FT misc_feature 405. .406
 FT /*tag= d
 FT /label= exon_junction
 FT misc_feature 484. .546
 FT /*tag= x
 FT /label= membrane-spanning_segment
 FT misc_feature 621. .622
 FT /*tag= e
 FT /label= exon_junction
 FT misc_feature 711. .712
 FT /*tag= f
 FT /label= exon_junction
 FT misc_feature 714. .777
 FT /*tag= y
 FT /label= membrane-spanning_segment
 FT misc_feature 793. .855
 FT /*tag= z
 FT /label= membrane-spanning_segment
 FT misc_feature 1001. .1002
 FT /*tag= g
 FT /label= exon_junction
 FT misc_feature 1054. .1116
 FT /*tag= aa
 FT /label= membrane-spanning_segment
 FT misc_feature 1120. .1182
 FT /*tag= ab
 FT /label= membrane-spanning_segment
 FT misc_feature 1248. .1249
 FT /*tag= h
 FT /label= exon_junction
 FT misc_feature 1341. .1342
 FT /*tag= i

FT		/label= exon_junction
FT	misc_binding	1429. .1884
FT		/*tag= ai
FT		/label= ATP-binding_fold
FT	misc_feature	1523. .1524
FT		/*tag= j
FT		/label= exon_junction
FT	misc_feature	1716. .1717
FT		/*tag= k
FT		/label= exon_junction
FT	misc_feature	1811. .1812
FT		/*tag= l
FT		/label= exon_junction
FT	misc_feature	1898. .1899
FT		/*tag= m
FT		/label= exon_junction
FT	misc_feature	2622. .2623
FT		/*tag= n
FT		/label= exon_junction
FT	misc_feature	2710. .2772
FT		/*tag= ac
FT		/label= membrane-spanning_segment
FT	misc_feature	2789. .2790
FT		/*tag= o
FT		/label= exon_junction
FT	misc_feature	2866. .2928
FT		/*tag= ad
FT		/label= membrane-spanning_segment
FT	misc_feature	3040. .3041
FT		/*tag= p
FT		/label= exon_junction
FT	misc_feature	3103. .3165
FT		/*tag= ae
FT		/label= membrane-spanning_segment
FT	misc_feature	3171. .3234
FT		/*tag= af
FT		/label= membrane-spanning_segment
FT	misc_feature	3439. .3501
FT		/*tag= ag
FT		/label= membrane-spanning_segment
FT	misc_feature	3499. .3500
FT		/*tag= q
FT		/label= exon_junction
FT	misc_feature	3517. .3582
FT		/*tag= ah
FT		/label= membrane-spanning_segment
FT	misc_feature	3599. .3600
FT		/*tag= r
FT		/label= exon_junction
FT	misc_binding	3787. .4290
FT		/*tag= aj
FT		/label= ATP-binding_fold
FT	misc_feature	3849. .3850
FT		/*tag= s
FT		/label= exon_junction
FT	misc_feature	4095. .4096
FT		/*tag= t

FT /label= exon_junction
 FT misc_feature 4268. .4269
 FT /*tag= u
 FT /label= exon_junction
 FT misc_feature 4374. .4375
 FT /*tag= v
 FT /label= exon_junction

XX
 PN WO9110734-A.

XX
 PD 25-JUL-1991.

XX
 PF 12-JAN-1990; 90CA-02007699.

XX
 PR 12-JAN-1990; 90CA-02007699.

PR 01-MAR-1990; 90CA-02011253.

PR 10-JUL-1990; 90CA-02020817.

XX
 PA (HSCR-) HSC RES DEV CORP.

XX
 PI Tsui LC, Rommens JM, Kerem B;

XX
 DR WPI; 1991-238022/32.

DR P-PSDB; AAR13300.
 XX
 PT Mutant cystic fibrosis trans-membrane conductance regulator gene - used
 PT for producing prods. for diagnosis, screening and therapy of cystic
 PT fibrosis.

XX
 PS Claim 1; Page 121; 178pp; English.

XX
 CC In the Y563N mutation a T to A change is detected at nucleotide position
 CC 1820 in exon 12. The mutant CF gene when expressed in cells of the human
 CC body, is associated with altered cell function which correlates with the
 CC genetic disease cystic fibrosis. See also AAQ13053-72

XX
 SQ Sequence 6129 BP; 1886 A; 1183 C; 1329 G; 1731 T; 0 U; 0 Other;

Query Match 31.6%; Score 32.2; DB 2; Length 6129;
 Best Local Similarity 63.6%; Pred. No. 0.83;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 5 TAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || ||| ||||| || | | | ||| || | ||| | || |||
 Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 1604
 QY 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 1605 TTCTCAGTTTTCTGGA 1621

Search completed: April 29, 2004, 15:06:51
 Job time : 53.1699 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 29, 2004, 14:53:14 ; Search time 9.33373 Seconds
(without alignments)
6064.561 Million cell updates/sec

Title: US-09-989-981A-9_COPY_3_104
Perfect score: 102
Sequence: 1 ctggtaggtgagatctctga.....aacaagctgtcctggaggcc 102

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		%		Query			ID	Description
Result	No.	Score	Match	Length	DB			
c	1	32.2	31.6	420	3	US-09-158-863C-64	Sequence 64, Appl	
	2	32.2	31.6	551	2	US-08-647-368A-4	Sequence 4, Appli	
	3	32.2	31.6	558	2	US-08-647-368A-3	Sequence 3, Appli	
	4	32.2	31.6	795	4	US-09-866-293-9	Sequence 9, Appli	
	5	32.2	31.6	2640	1	US-08-216-971-1	Sequence 1, Appli	
	6	32.2	31.6	2640	2	US-08-812-979-1	Sequence 1, Appli	
	7	32.2	31.6	2908	3	US-08-487-799-1	Sequence 1, Appli	
	8	32.2	31.6	4443	4	US-09-425-453A-1	Sequence 1, Appli	
	9	32.2	31.6	4443	4	US-09-425-453A-3	Sequence 3, Appli	
	10	32.2	31.6	4443	4	US-09-425-453A-5	Sequence 5, Appli	
	11	32.2	31.6	4443	4	US-09-425-453A-7	Sequence 7, Appli	

12	32.2	31.6	4443	4	US-09-425-453A-9	Sequence 9, Appli
13	32.2	31.6	4443	4	US-09-425-453A-11	Sequence 11, Appl
14	32.2	31.6	4443	4	US-09-425-453A-13	Sequence 13, Appl
15	32.2	31.6	4443	4	US-09-425-453A-15	Sequence 15, Appl
16	32.2	31.6	4443	4	US-09-425-453A-17	Sequence 17, Appl
17	32.2	31.6	4443	4	US-09-425-453A-19	Sequence 19, Appl
18	32.2	31.6	4560	3	US-09-256-703-1	Sequence 1, Appli
19	32.2	31.6	5635	1	US-08-136-742A-3	Sequence 3, Appli
20	32.2	31.6	5635	3	US-09-248-026-3	Sequence 3, Appli
21	32.2	31.6	5635	5	PCT-US93-11667-3	Sequence 3, Appli
22	32.2	31.6	6126	2	US-08-951-912-3	Sequence 3, Appli
23	32.2	31.6	6126	4	US-09-174-077-3	Sequence 3, Appli
24	32.2	31.6	6129	1	US-07-637-621-1	Sequence 1, Appli
25	32.2	31.6	6129	1	US-08-136-742A-1	Sequence 1, Appli
26	32.2	31.6	6129	1	US-08-135-809A-1	Sequence 1, Appli
27	32.2	31.6	6129	2	US-08-951-912-1	Sequence 1, Appli
28	32.2	31.6	6129	2	US-08-951-912-5	Sequence 5, Appli
29	32.2	31.6	6129	2	US-08-691-605-1	Sequence 1, Appli
30	32.2	31.6	6129	3	US-09-248-026-1	Sequence 1, Appli
31	32.2	31.6	6129	3	US-08-681-838A-1	Sequence 1, Appli
32	32.2	31.6	6129	4	US-09-174-077-1	Sequence 1, Appli
33	32.2	31.6	6129	4	US-09-174-077-5	Sequence 5, Appli
34	32.2	31.6	6129	5	PCT-US93-11667-1	Sequence 1, Appli
35	32.2	31.6	6130	1	US-08-466-886-16	Sequence 16, Appl
36	32.2	31.6	6130	2	US-08-604-488-1	Sequence 1, Appli
37	32.2	31.6	6130	2	US-08-469-461-1	Sequence 1, Appli
38	32.2	31.6	6130	3	US-07-890-609-1	Sequence 1, Appli
39	32.2	31.6	6130	3	US-08-030-081-1	Sequence 1, Appli
40	32.2	31.6	6130	3	US-08-469-617-16	Sequence 16, Appl
41	32.2	31.6	6146	6	5240846-4	Patent No. 5240846
42	32.2	31.6	8225	3	US-08-793-618-1	Sequence 1, Appli
43	32.2	31.6	8225	4	US-09-794-431-1	Sequence 1, Appli
c 44	32.2	31.6	9972	3	US-08-836-022A-3	Sequence 3, Appli
c 45	32.2	31.6	9972	3	US-09-427-048A-3	Sequence 3, Appli
46	32.2	31.6	12143	4	US-09-423-744A-1	Sequence 1, Appli
47	32.2	31.6	22846	2	US-08-469-461-3	Sequence 3, Appli
48	32.2	31.6	22846	3	US-07-890-609-3	Sequence 3, Appli
c 49	30.2	29.6	1251	4	US-09-252-991A-1019	Sequence 1019, Ap
50	30.2	29.6	2847	4	US-09-252-991A-1036	Sequence 1036, Ap

ALIGNMENTS

RESULT 1

US-09-158-863C-64

; Sequence 64, Application US/09158863C

; Patent No. 6280978

; GENERAL INFORMATION:

; APPLICANT: Mitchell, Lloyd G.

; APPLICANT: Garcia-Blanco, Mariano A.

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR USE IN

; TITLE OF INVENTION: SPLICEOSOME MEDIATED RNA TRANS-SPLICING

; FILE REFERENCE: 31304-B-A

; CURRENT APPLICATION NUMBER: US/09/158,863C

; CURRENT FILING DATE: 1998-09-23

; PRIOR APPLICATION NUMBER: 09/133,717

```
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 09/087,233
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 08/766,354
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: 60/008,317
; PRIOR FILING DATE: 1995-12-07
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 64
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: trans-spliced product comprising cystic fibrosis
; OTHER INFORMATION: transmembrane regulator-derived sequences and His
; OTHER INFORMATION: tag sequences
US-09-158-863C-64
```

```
Query Match          31.6%; Score 32.2; DB 3; Length 420;
Best Local Similarity 63.6%; Pred. No. 0.02;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
```

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
        || | |||| ||   ||| ||||| ||   | | | ||| || |   ||| | || -|||
Db     128 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTCATTCTG 187

Qy      65 TTGTCACTTTCCGAGGA 81
        || ||| ||| |   |||
Db     188 TTCTCAGTTTTCTGGA 204
```

RESULT 2

US-08-647-368A-4

```
; Sequence 4, Application US/08647368A
; Patent No. 5928906
; GENERAL INFORMATION:
; APPLICANT: Koster, Hubert
; APPLICANT: Van de Boom, Dirk
; APPLICANT: Ruppert, Andreas
; TITLE OF INVENTION: PROCESS FOR DIRECT SEQUENCING DURING
; TITLE OF INVENTION: TEMPLATE AMPLIFICATION
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/647,368A
```

```

; FILING DATE: 09-MAY-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: SQA-020.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 551 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-647-368A-4

```

```

Query Match          31.6%; Score 32.2; DB 2; Length 551;
Best Local Similarity 63.6%; Pred. No. 0.023;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
        || | ||| ||   || | ||| ||   | | | ||| || |   || | ||| |||
Db      85 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTCATTCTG 144

Qy      65 TTGTCACTTTCCGAGGA 81
        || ||| ||| |   |||
Db      145 TTCTCAGTTTTCCTGGA 161

```

RESULT 3

US-08-647-368A-3/c

; Sequence 3, Application US/08647368A

; Patent No. 5928906

; GENERAL INFORMATION:

; APPLICANT: Koster, Hubert

; APPLICANT: Van de Boom, Dirk

; APPLICANT: Ruppert, Andreas

; TITLE OF INVENTION: PROCESS FOR DIRECT SEQUENCING DURING

; TITLE OF INVENTION: TEMPLATE AMPLIFICATION

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FOLEY, HOAG & ELIOT LLP

; STREET: One Post Office Square

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109-2170

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/647,368A

; FILING DATE: 09-MAY-1996


```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: SQA-020.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 558 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-647-368A-3

```

```

Query Match          31.6%; Score 32.2; DB 2; Length 558;
Best Local Similarity 63.6%; Pred. No. 0.023;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
        || | |||| ||   ||| |||| ||   | | | ||| || | ||| | || |||
Db      432 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTCATTCTG 373

Qy      65 TTGTCACTTTCCGAGGA 81
        || ||| ||| | |||
Db      372 TTCTCAGTTTTCCTGGA 356

```

RESULT 4

US-09-866-293-9

```

; Sequence 9, Application US/09866293
; Patent No. 6607911

```

GENERAL INFORMATION:

```

; APPLICANT: Gordon, Joan
; APPLICANT: Rundell, Clark
; TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATING TO CONTROL DNA
CONSTRUCT

```

```

; FILE REFERENCE: 053689-5010
; CURRENT APPLICATION NUMBER: US/09/866,293
; CURRENT FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 795
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-293-9

```

```

Query Match          31.6%; Score 32.2; DB 4; Length 795;
Best Local Similarity 63.6%; Pred. No. 0.026;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
        || | |||| ||   ||| |||| ||   | | | ||| || | ||| | || |||
Db      369 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTCATTCTG 428

```

Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 429 TTCTCAGTTTTCTGGA 445

RESULT 5

US-08-216-971-1

; Sequence 1, Application US/08216971
 ; Patent No. 5639661
 ; GENERAL INFORMATION:
 ; APPLICANT: Welsh, Michael J.
 ; APPLICANT: Sheppard, David N.
 ; TITLE OF INVENTION: NOVEL GENES AND PROTEINS FOR TREATING
 ; TITLE OF INVENTION: CYSTIC FIBROSIS
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 State Street, suite #510
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109-1875
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/216,971
 ; FILING DATE: 23-MAR-1994
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Arnold, Beth E.
 ; REGISTRATION NUMBER: 35,430
 ; REFERENCE/DOCKET NUMBER: UIZ-011
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 227-7400
 ; TELEFAX: (617) 227-5941
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2640 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 133..2640
 US-08-216-971-1

Query Match 31.6%; Score 32.2; DB 1; Length 2640;
 Best Local Similarity 63.6%; Pred. No. 0.042;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || ||| ||||| || | | | ||| || | ||| | || |||

Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAGAATTTCATTCTG 1604

Qy 65 TTGTCACTTCCGAGGA 81
 || ||| ||| | |||

Db 1605 TTCTCAGTTTCCTGGA 1621

RESULT 6

US-08-812-979-1

; Sequence 1, Application US/08812979
 ; Patent No. 5958893
 ; GENERAL INFORMATION:
 ; APPLICANT: Welsh, Michael J.
 ; APPLICANT: Sheppard, David N.
 ; TITLE OF INVENTION: NOVEL GENES AND PROTEINS FOR TREATING
 ; TITLE OF INVENTION: CYSTIC FIBROSIS
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 State Street, suite #510
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109-1875
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/812,979
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/216,971
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Arnold, Beth E.
 ; REGISTRATION NUMBER: 35,430
 ; REFERENCE/DOCKET NUMBER: UIZ-011
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 227-7400
 ; TELEFAX: (617) 227-5941
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2640 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 133..2640

US-08-812-979-1

Query Match 31.6%; Score 32.2; DB 2; Length 2640;
 Best Local Similarity 63.6%; Pred. No. 0.042;

Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
      || | ||| ||   ||| |||| ||   | | | ||| || | ||| | || |||
Db      1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 1604

Qy      65 TTGTCACCTTCCGAGGA 81
      || ||| ||| | |||
Db      1605 TTCTCAGTTTTCCTGGA 1621
```

RESULT 7

US-08-487-799-1

; Sequence 1, Application US/08487799C

; Patent No. 6010908

; GENERAL INFORMATION:

; APPLICANT: Gruenert, Deiter C.

; APPLICANT: Kunzelmann, Karl

; TITLE OF INVENTION: GENE THERAPY BY SMALL FRAGMENTS HOMOLOGOUS REPLACEMENT

; FILE REFERENCE: 480.18-1(HV)

; CURRENT APPLICATION NUMBER: US/08/487,799C

; CURRENT FILING DATE: 1995-06-07

; EARLIER APPLICATION NUMBER: 07/933,471

; EARLIER FILING DATE: 1992-08-21

; EARLIER APPLICATION NUMBER: 08/409,544

; EARLIER FILING DATE: 1995-03-24

; NUMBER OF SEQ ID NOS: 87

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 2908

; TYPE: DNA

; ORGANISM: human

US-08-487-799-1

Query Match 31.6%; Score 32.2; DB 3; Length 2908;

Best Local Similarity 63.6%; Pred. No. 0.044;

Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
      || | ||| ||   ||| |||| ||   | | | ||| || | ||| | || |||
Db      1085 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 1144

Qy      65 TTGTCACCTTCCGAGGA 81
      || ||| ||| | |||
Db      1145 TTCTCAGTTTTCCTGGA 1161
```

RESULT 8

US-09-425-453A-1

; Sequence 1, Application US/09425453A

; Patent No. 6468793

; GENERAL INFORMATION:

; APPLICANT: Teem, John L.

; TITLE OF INVENTION: CFTR Genes and Proteins for Cystic Fibrosis Gene Therapy

; FILE REFERENCE: FSU-99XC1

; CURRENT APPLICATION NUMBER: US/09/425,453A

; CURRENT FILING DATE: 1999-10-22

; PRIOR APPLICATION NUMBER: 60/105,444
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4443
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(4443)
US-09-425-453A-1

Query Match 31.6%; Score 32.2; DB 4; Length 4443;
Best Local Similarity 63.6%; Pred. No. 0.052;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| |||| || | | | ||| || | ||| | || |||
Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTCACTCTG 1472

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 1473 TTCTCAGTTTTCCTGGA 1489

RESULT 9

US-09-425-453A-3

; Sequence 3, Application US/09425453A
; Patent No. 6468793
; GENERAL INFORMATION:
; APPLICANT: Teem, John L.
; TITLE OF INVENTION: CFTR Genes and Proteins for Cystic Fibrosis Gene Therapy
; FILE REFERENCE: FSU-99XC1
; CURRENT APPLICATION NUMBER: US/09/425,453A
; CURRENT FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,444
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 4443
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(4443)
US-09-425-453A-3

Query Match 31.6%; Score 32.2; DB 4; Length 4443;
Best Local Similarity 63.6%; Pred. No. 0.052;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| |||| || | | | ||| || | ||| | || |||
Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTCACTCTG 1472

Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 1473 TTCTCAGTTTTCCTGGA 1489

RESULT 10

US-09-425-453A-5
 ; Sequence 5, Application US/09425453A
 ; Patent No. 6468793
 ; GENERAL INFORMATION:
 ; APPLICANT: Teem, John L.
 ; TITLE OF INVENTION: CFTR Genes and Proteins for Cystic Fibrosis Gene Therapy
 ; FILE REFERENCE: FSU-99XC1
 ; CURRENT APPLICATION NUMBER: US/09/425,453A
 ; CURRENT FILING DATE: 1999-10-22
 ; PRIOR APPLICATION NUMBER: 60/105,444
 ; PRIOR FILING DATE: 1998-10-23
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 4443
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-425-453A-5

Query Match 31.6%; Score 32.2; DB 4; Length 4443;
 Best Local Similarity 63.6%; Pred. No. 0.052;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || ||| ||||| || | | | ||| || | ||| | || |||
 Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTCATTCTG 1472

Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 1473 TTCTCAGTTTTCCTGGA 1489

RESULT 11

US-09-425-453A-7
 ; Sequence 7, Application US/09425453A
 ; Patent No. 6468793
 ; GENERAL INFORMATION:
 ; APPLICANT: Teem, John L.
 ; TITLE OF INVENTION: CFTR Genes and Proteins for Cystic Fibrosis Gene Therapy
 ; FILE REFERENCE: FSU-99XC1
 ; CURRENT APPLICATION NUMBER: US/09/425,453A
 ; CURRENT FILING DATE: 1999-10-22
 ; PRIOR APPLICATION NUMBER: 60/105,444
 ; PRIOR FILING DATE: 1998-10-23
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 7
 ; LENGTH: 4443
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:

; NAME/KEY: gene
; LOCATION: (1)..(4443)
US-09-425-453A-7

Query Match 31.6%; Score 32.2; DB 4; Length 4443;
Best Local Similarity 63.6%; Pred. No. 0.052;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
QY          5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
              || | ||| ||   || | ||| ||   | | | ||| || | ||| | || | |||
Db         1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 1472

QY          65 TTGTCACTTTCCGAGGA 81
              || ||| ||| | |||
Db         1473 TTCTCAGTTTTCTGGA 1489
```

RESULT 12

US-09-425-453A-9

; Sequence 9, Application US/09425453A
; Patent No. 6468793
; GENERAL INFORMATION:
; APPLICANT: Teem, John L.
; TITLE OF INVENTION: CFTR Genes and Proteins for Cystic Fibrosis Gene Therapy
; FILE REFERENCE: FSU-99XC1
; CURRENT APPLICATION NUMBER: US/09/425,453A
; CURRENT FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,444
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 4443
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(4443)
US-09-425-453A-9

Query Match 31.6%; Score 32.2; DB 4; Length 4443;
Best Local Similarity 63.6%; Pred. No. 0.052;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
QY          5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
              || | ||| ||   || | ||| ||   | | | ||| || | ||| | || | |||
Db         1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 1472

QY          65 TTGTCACTTTCCGAGGA 81
              || ||| ||| | |||
Db         1473 TTCTCAGTTTTCTGGA 1489
```

RESULT 13

US-09-425-453A-11

; Sequence 11, Application US/09425453A
; Patent No. 6468793

; GENERAL INFORMATION:
; APPLICANT: Teem, John L.
; TITLE OF INVENTION: CFTR Genes and Proteins for Cystic Fibrosis Gene Therapy
; FILE REFERENCE: FSU-99XC1
; CURRENT APPLICATION NUMBER: US/09/425,453A
; CURRENT FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,444
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 4443
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-425-453A-11

Query Match 31.6%; Score 32.2; DB 4; Length 4443;
Best Local Similarity 63.6%; Pred. No. 0.052;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
||| ||||| || ||| ||||| || | | | ||| || | ||| ||| |||
Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTCATTCTG 1472

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 1473 TTCTCAGTTTTCTCTGGA 1489

RESULT 14

US-09-425-453A-13

; Sequence 13, Application US/09425453A
; Patent No. 6468793
; GENERAL INFORMATION:
; APPLICANT: Teem, John L.
; TITLE OF INVENTION: CFTR Genes and Proteins for Cystic Fibrosis Gene Therapy
; FILE REFERENCE: FSU-99XC1
; CURRENT APPLICATION NUMBER: US/09/425,453A
; CURRENT FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,444
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 4443
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-425-453A-13

Query Match 31.6%; Score 32.2; DB 4; Length 4443;
Best Local Similarity 63.6%; Pred. No. 0.052;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
||| ||||| || ||| ||||| || | | | ||| || | ||| ||| |||
Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTCATTCTG 1472

Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 1473 TTCTCAGTTTTCCTGGA 1489

RESULT 15

US-09-425-453A-15
 ; Sequence 15, Application US/09425453A
 ; Patent No. 6468793
 ; GENERAL INFORMATION:
 ; APPLICANT: Teem, John L.
 ; TITLE OF INVENTION: CFTR Genes and Proteins for Cystic Fibrosis Gene Therapy
 ; FILE REFERENCE: FSU-99XC1
 ; CURRENT APPLICATION NUMBER: US/09/425,453A
 ; CURRENT FILING DATE: 1999-10-22
 ; PRIOR APPLICATION NUMBER: 60/105,444
 ; PRIOR FILING DATE: 1998-10-23
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 15
 ; LENGTH: 4443
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: gene
 ; LOCATION: (1)..(4443)
 US-09-425-453A-15

Query Match 31.6%; Score 32.2; DB 4; Length 4443;
 Best Local Similarity 63.6%; Pred. No. 0.052;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || ||| ||||| || | | | ||| || | ||| | || |||
 Db 1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTTTCATTCTG 1472

Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 1473 TTCTCAGTTTTCCTGGA 1489

RESULT 16

US-09-425-453A-17
 ; Sequence 17, Application US/09425453A
 ; Patent No. 6468793
 ; GENERAL INFORMATION:
 ; APPLICANT: Teem, John L.
 ; TITLE OF INVENTION: CFTR Genes and Proteins for Cystic Fibrosis Gene Therapy
 ; FILE REFERENCE: FSU-99XC1
 ; CURRENT APPLICATION NUMBER: US/09/425,453A
 ; CURRENT FILING DATE: 1999-10-22
 ; PRIOR APPLICATION NUMBER: 60/105,444
 ; PRIOR FILING DATE: 1998-10-23
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 17
 ; LENGTH: 4443

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-425-453A-17

Query Match 31.6%; Score 32.2; DB 4; Length 4443;
Best Local Similarity 63.6%; Pred. No. 0.052;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
          || | |||| ||   ||| ||||| ||   | | | ||| || | ||| | || |||
Db      1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTCATTCTG 1472

Qy      65 TTGTCACTTTCCGAGGA 81
          || ||| ||| | |||
Db      1473 TTCTCAGTTTTCCTGGA 1489
```

RESULT 17

US-09-425-453A-19
; Sequence 19, Application US/09425453A
; Patent No. 6468793
; GENERAL INFORMATION:
; APPLICANT: Teem, John L.
; TITLE OF INVENTION: CFTR Genes and Proteins for Cystic Fibrosis Gene Therapy
; FILE REFERENCE: FSU-99XC1
; CURRENT APPLICATION NUMBER: US/09/425,453A
; CURRENT FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,444
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 4443
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-425-453A-19

Query Match 31.6%; Score 32.2; DB 4; Length 4443;
Best Local Similarity 63.6%; Pred. No. 0.052;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
          || | |||| ||   ||| ||||| ||   | | | ||| || | ||| | || |||
Db      1413 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTCATTCTG 1472

Qy      65 TTGTCACTTTCCGAGGA 81
          || ||| ||| | |||
Db      1473 TTCTCAGTTTTCCTGGA 1489
```

RESULT 18

US-09-256-703-1
; Sequence 1, Application US/09256703
; Patent No. 6294379
; GENERAL INFORMATION:
; APPLICANT: Dong, Jian-yun
; APPLICANT: Kan, Yuet Wai

```
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Efficient AAV Vectors
; FILE REFERENCE: 023070-084910US
; CURRENT APPLICATION NUMBER: US/09/256,703
; CURRENT FILING DATE: 1999-02-24
; PRIOR APPLICATION NUMBER: US 60/075,980
; PRIOR FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4560
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: truncated cystic fibrosis transmembrane
; OTHER INFORMATION: conductance regulator (CFTR) polynucleotide
; OTHER INFORMATION: encoding a functional CFTR polypeptide
; NAME/KEY: CDS
; LOCATION: (133)..(4560)
US-09-256-703-1
```

```
Query Match          31.6%; Score 32.2; DB 3; Length 4560;
Best Local Similarity 63.6%; Pred. No. 0.052;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
```

```
Qy          5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
      || | ||| ||      ||| ||||| || | | | ||| || | ||| ||| |||
Db          1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTCAATTCTG 1604

Qy          65 TTGTCACTTTCCGAGGA 81
      || ||| ||| | |||
Db          1605 TTCTCAGTTTTCCTGGA 1621
```

RESULT 19

US-08-136-742A-3

```
; Sequence 3, Application US/08136742A
; Patent No. 5670488
; GENERAL INFORMATION:
; APPLICANT: Gregory, R.J., Armentano, D., Couture, L.A., Smith,
; APPLICANT: A.E.
; TITLE OF INVENTION: GENE THERAPY FOR CYSTIC FIBROSIS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRUMBAUGH, GRAVES, DONOHUE & RAYMOND
; STREET: 30 ROCKEFELLER PLAZA
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/136,742A
```

```

; FILING DATE: 02-DEC-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/985,478
; FILING DATE: 02-DEC-1992
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Seide, Rochelle K.
; REGISTRATION NUMBER: 32,300
; REFERENCE/DOCKET NUMBER: A30668 (Genzyme Dkt. IG4-9.11)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 408-2500
; TELEFAX: (212) 765-2519
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5635 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-136-742A-3

```

```

Query Match          31.6%; Score 32.2; DB 1; Length 5635;
Best Local Similarity 63.6%; Pred. No. 0.057;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
      || | |||| || ||| ||||| || | | ||| || | ||| || |||
Db      2015 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTTCATTCTG 2074

Qy      65 TTGTCACTTTCCGAGGA 81
      || ||| ||| | |||
Db      2075 TTCTCAGTTTTCCTGGA 2091

```

RESULT 20

US-09-248-026-3

```

; Sequence 3, Application US/09248026
; Patent No. 6093567
; GENERAL INFORMATION:
; APPLICANT: Gregory, R.J., Armentano, D., Couture, L.A., Smith,
; APPLICANT: A.E.
; TITLE OF INVENTION: ADENOVIRUS VECTORS FOR GENE THERAPY
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTTS, L.L.P.
; STREET: 30 ROCKEFELLER PLAZA
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/09/248,026
; FILING DATE: 10-FEB-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/895,194
; FILING DATE: 16-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Seide, Rochelle K.
; REGISTRATION NUMBER: 32,300
; REFERENCE/DOCKET NUMBER: A30668-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 705-5000
; TELEFAX: (212) 705-5020
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5635 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-248-026-3

```

```

Query Match          31.6%; Score 32.2; DB 3; Length 5635;
Best Local Similarity 63.6%; Pred. No. 0.057;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
      || | ||| || ||| ||||| || | | ||| || | ||| ||| |||
Db      2015 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG 2074

Qy      65 TTGTCACTTTCCGAGGA 81
      || ||| ||| | |||
Db      2075 TTCTCAGTTTTTCCTGGA 2091

```

RESULT 21

PCT-US93-11667-3

```

; Sequence 3, Application PC/TUS9311667
; GENERAL INFORMATION:
; APPLICANT: Gregory, R.J., Armentano, D., Couture, L.A., Smith,
; APPLICANT: A.E.
; TITLE OF INVENTION: GENE THERAPY FOR CYSTIC FIBROSIS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11667

```

```
; FILING DATE: 02-DEC-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/985,478
; FILING DATE: 02-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanley, Elizabeth A.
; REGISTRATION NUMBER: 33,505
; REFERENCE/DOCKET NUMBER: NZI-014CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5635 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
PCT-US93-11667-3
```

```
Query Match          31.6%; Score 32.2; DB 5; Length 5635;
Best Local Similarity 63.6%; Pred. No. 0.057;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
```

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
        || | |||| ||      ||| ||||| ||      | | | ||| || | ||| |||
Db      2015 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGGAGAATTTCATTCTG 2074

Qy      65 TTGTCACTTTCCGAGGA 81
        || ||| ||| | |||
Db      2075 TTCTCAGTTTTCCTGGA 2091
```

RESULT 22

US-08-951-912-3

```
; Sequence 3, Application US/08951912
; Patent No. 5972995
; GENERAL INFORMATION:
; APPLICANT: Fischer, Horst
; APPLICANT: Illek, Beate
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CYSTIC
; TITLE OF INVENTION: FIBROSIS THERAPY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```

; APPLICATION NUMBER: US/08/951,912
; FILING DATE: 16-OCT-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 200116.403
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6126 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 133..4569
US-08-951-912-3

```

```

Query Match          31.6%; Score 32.2; DB 2; Length 6126;
Best Local Similarity 63.6%; Pred. No. 0.059;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
        || | |||| ||      || |||| || | | | ||| || | ||| || | |||
Db      1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTCACTCTG 1604

Qy      65 TTGTCACTTTCCGAGGA 81
        || ||| ||| | |||
Db      1605 TTCTCAGTTTTCTGGA 1621

```

RESULT 23

```

US-09-174-077-3
; Sequence 3, Application US/09174077
; Patent No. 6329422
; GENERAL INFORMATION:
; APPLICANT: Fischer, Horst
; APPLICANT: Illek, Beate
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CYSTIC FIBROSIS THERAPY
; FILE REFERENCE: 200116.403C1
; CURRENT APPLICATION NUMBER: US/09/174,077
; CURRENT FILING DATE: 1998-10-16
; EARLIER APPLICATION NUMBER: US 08/951,912
; EARLIER FILING DATE: 1997-10-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 6126
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-174-077-3

```

```

Query Match          31.6%; Score 32.2; DB 4; Length 6126;
Best Local Similarity 63.6%; Pred. No. 0.059;

```

Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || ||| |||| || | | | ||| || | ||| | || |||
 Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTCATTCTG 1604

Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 1605 TTCTCAGTTTTCTGGA 1621

RESULT 24

US-07-637-621-1

; Sequence 1, Application US/07637621
 ; Patent No. 5407796
 ; GENERAL INFORMATION:
 ; APPLICANT: cutting, gary
 ; APPLICANT: antonarakis, stylianos e
 ; APPLICANT: kazazian jr., haig h
 ; TITLE OF INVENTION: CYSTIC FIBROSIS MUTATION CLUSTER
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Banner, Birch, McKie and Beckett
 ; STREET: 1001 G Street, N.W.
 ; CITY: Washington, D.C.
 ; COUNTRY: USA
 ; ZIP: 20001
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/637,621
 ; FILING DATE: 19910104
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: kagan, sarah a
 ; REGISTRATION NUMBER: 32,141
 ; REFERENCE/DOCKET NUMBER: 1107.030010
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-508-9100
 ; TELEFAX: 202-508-9100
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 6129 base pairs
 ; TYPE: NUCLEIC ACID
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Homo sapiens
 US-07-637-621-1

Query Match 31.6%; Score 32.2; DB 1; Length 6129;

Best Local Similarity 63.6%; Pred. No. 0.059;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy          5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
             || | |||| ||      ||| ||||| ||  | | | ||| || | ||| | || |||
Db        1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTTCATTCTG 1604

Qy          65 TTGTCACTTTCCGAGGA 81
             || ||| ||| | |||
Db        1605 TTCTCAGTTTTCCTGGA 1621
```

RESULT 25

US-08-136-742A-1

```
; Sequence 1, Application US/08136742A
; Patent No. 5670488
; GENERAL INFORMATION:
;   APPLICANT: Gregory, R.J., Armentano, D., Couture, L.A., Smith,
;   APPLICANT: A.E.
;   TITLE OF INVENTION: GENE THERAPY FOR CYSTIC FIBROSIS
;   NUMBER OF SEQUENCES: 10
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: BRUMBAUGH, GRAVES, DONOHUE & RAYMOND
;     STREET: 30 ROCKEFELLER PLAZA
;     CITY: NEW YORK
;     STATE: NEW YORK
;     COUNTRY: USA
;     ZIP: 10112
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: ASCII
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/136,742A
;     FILING DATE: 02-DEC-1993
;     CLASSIFICATION: 514
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 07/985,478
;     FILING DATE: 02-DEC-1992
;     CLASSIFICATION: 514
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Seide, Rochelle K.
;     REGISTRATION NUMBER: 32,300
;     REFERENCE/DOCKET NUMBER: A30668 (Genzyme Dkt. IG4-9.11)
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (212) 408-2500
;     TELEFAX: (212) 765-2519
;   INFORMATION FOR SEQ ID NO: 1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 6129 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;     MOLECULE TYPE: cDNA
;     FEATURE:
;       NAME/KEY: CDS
```

; LOCATION: 133..4572
US-08-136-742A-1

Query Match 31.6%; Score 32.2; DB 1; Length 6129;
Best Local Similarity 63.6%; Pred. No. 0.059;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| |||| || | | | || | || | ||
Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 1604

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 1605 TTCTCAGTTTTCTGGA 1621

RESULT 26

US-08-135-809A-1

; Sequence 1, Application US/08135809A

; Patent No. 5688677

; GENERAL INFORMATION:

; APPLICANT: CHENG, SENG H.

; APPLICANT: DITULLIO, PAUL

; APPLICANT: EBERT, KARL M.

; APPLICANT: MEADE, HARRY M.

; APPLICANT: SMITH, ALAN E.

; TITLE OF INVENTION: DEOXYRIBONUCLEIC ACIDS CONTAINING

; TITLE OF INVENTION: INACTIVATED HORMONE RESPONSIVE ELEMENTS

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENZYME CORPORATION

; STREET: ONE MOUNTAIN ROAD

; CITY: FRAMINGHAM

; STATE: MASSACHUSETTS

; COUNTRY: USA

; ZIP: 01701

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/135,809A

; FILING DATE: 13-OCT-1993

; CLASSIFICATION: 800

; ATTORNEY/AGENT INFORMATION:

; NAME: LASSEN, ELIZABETH

; REGISTRATION NUMBER: 31,845

; REFERENCE/DOCKET NUMBER: IG4-9.12

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (508) 872-8400

; TELEFAX: (508) 872-5415

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 6129 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 133..4572
US-08-135-809A-1

Query Match 31.6%; Score 32.2; DB 1; Length 6129;
Best Local Similarity 63.6%; Pred. No. 0.059;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
||| ||||| || ||| ||||| || | | | ||| || | ||| | || |||
Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG 1604

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 1605 TTCTCAGTTTTCTGGA 1621

RESULT 27

US-08-951-912-1

; Sequence 1, Application US/08951912
; Patent No. 5972995
; GENERAL INFORMATION:
; APPLICANT: Fischer, Horst
; APPLICANT: Illek, Beate
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CYSTIC
; TITLE OF INVENTION: FIBROSIS THERAPY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/951,912
; FILING DATE: 16-OCT-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 200116.403
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6129 base pairs
; TYPE: nucleic acid

; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 133..4572
US-08-951-912-1

Query Match 31.6%; Score 32.2; DB 2; Length 6129;
Best Local Similarity 63.6%; Pred. No. 0.059;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| |||| || | | | ||| || | ||| ||| |||
Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 1604

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 1605 TTCTCAGTTTTCCTGGA 1621

RESULT 28

US-08-951-912-5

; Sequence 5, Application US/08951912
; Patent No. 5972995
; GENERAL INFORMATION:
; APPLICANT: Fischer, Horst
; APPLICANT: Illek, Beate
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CYSTIC
; TITLE OF INVENTION: FIBROSIS THERAPY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/951,912
; FILING DATE: 16-OCT-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 200116.403
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6129 base pairs
; TYPE: nucleic acid

; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-951-912-5

Query Match 31.6%; Score 32.2; DB 2; Length 6129;
Best Local Similarity 63.6%; Pred. No. 0.059;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| |||| || | | | || | | || | || ||
Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTCATTCTG 1604

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 1605 TTCTCAGTTTTCCTGGA 1621

RESULT 29

US-08-691-605-1

; Sequence 1, Application US/08691605
; Patent No. 5981714

; GENERAL INFORMATION:

; APPLICANT: Cheng, Seng H., Marshall, John, Gregory, Richard J.
; APPLICANT: and Rafter, Patrick. W.
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR CYSTIC FIBROSIS
; TITLE OF INVENTION: TRANSMEMBRANE CONDUCTANCE REGULATOR AND USES
; TITLE OF INVENTION: THEREFOR
; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/691,605
; FILING DATE:
; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/114,950
; FILING DATE:
; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Hanley, Elizabeth A.
; REGISTRATION NUMBER: 33,505
; REFERENCE/DOCKET NUMBER: NZI-029

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 6129 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 133..4572
US-08-691-605-1

Query Match 31.6%; Score 32.2; DB 2; Length 6129;
Best Local Similarity 63.6%; Pred. No. 0.059;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | ||| || ||| |||| || | | | ||| || | ||| | || |||
Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTTTCATTCTG 1604
Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 1605 TTCTCAGTTTTTCCTGGA 1621

RESULT 30

US-09-248-026-1

; Sequence 1, Application US/09248026

; Patent No. 6093567

; GENERAL INFORMATION:

; APPLICANT: Gregory, R.J., Armentano, D., Couture, L.A., Smith,

; APPLICANT: A.E.

; TITLE OF INVENTION: ADENOVIRUS VECTORS FOR GENE THERAPY

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BAKER & BOTTS, L.L.P.

; STREET: 30 ROCKEFELLER PLAZA

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10112

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/248,026

; FILING DATE: 10-FEB-1999

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/895,194

; FILING DATE: 16-JUL-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Seide, Rochelle K.

; REGISTRATION NUMBER: 32,300

; REFERENCE/DOCKET NUMBER: A30668-C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 705-5000

; TELEFAX: (212) 705-5020
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6129 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 133..4572
US-09-248-026-1

Query Match 31.6%; Score 32.2; DB 3; Length 6129;
Best Local Similarity 63.6%; Pred. No. 0.059;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
||| |||| || ||| |||| || | | | ||| || | ||| | || |||
Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTTTCATTCTG 1604

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 1605 TTCTCAGTTTTCCTGGA 1621

RESULT 31

US-08-681-838A-1

; Sequence 1, Application US/08681838A
; Patent No. 6245735
; GENERAL INFORMATION:
; APPLICANT: Pier, Gerald B
; TITLE OF INVENTION: Methods and Products for Treating
; TITLE OF INVENTION: Pseudomonas Infection
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks PC
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/681,838A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gates, Edward R
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: B0801/7054
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500

```

; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6129 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 133..4575
US-08-681-838A-1

```

```

Query Match          31.6%; Score 32.2; DB 3; Length 6129;
Best Local Similarity 63.6%; Pred. No. 0.059;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy          5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
             || | |||| ||      ||| ||||| ||      | | | ||| || | ||| | || |||
Db          1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTTCATTCTG 1604

Qy          65 TTGTCACCTTTCCGAGGA 81
             || ||| ||| | |||
Db          1605 TTCTCAGTTTTCTGGA 1621

```

RESULT 32

```

US-09-174-077-1
; Sequence 1, Application US/09174077
; Patent No. 6329422
; GENERAL INFORMATION:
; APPLICANT: Fischer, Horst
; APPLICANT: Illek, Beate
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CYSTIC FIBROSIS THERAPY
; FILE REFERENCE: 200116.403C1
; CURRENT APPLICATION NUMBER: US/09/174,077
; CURRENT FILING DATE: 1998-10-16
; EARLIER APPLICATION NUMBER: US 08/951,912
; EARLIER FILING DATE: 1997-10-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 6129
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-174-077-1

```

```

Query Match          31.6%; Score 32.2; DB 4; Length 6129;
Best Local Similarity 63.6%; Pred. No. 0.059;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy          5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
             || | |||| ||      ||| ||||| ||      | | | ||| || | ||| | || |||

```


Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTTTCATTCTG 1604

Qy 65 TTGTCACCTTTCCGAGGA 81
 || ||| ||| | |||

Db 1605 TTCTCAGTTTTTCCTGGA 1621

RESULT 33

US-09-174-077-5

; Sequence 5, Application US/09174077
 ; Patent No. 6329422
 ; GENERAL INFORMATION:
 ; APPLICANT: Fischer, Horst
 ; APPLICANT: Illek, Beate
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CYSTIC FIBROSIS THERAPY
 ; FILE REFERENCE: 200116.403C1
 ; CURRENT APPLICATION NUMBER: US/09/174,077
 ; CURRENT FILING DATE: 1998-10-16
 ; EARLIER APPLICATION NUMBER: US 08/951,912
 ; EARLIER FILING DATE: 1997-10-16
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 6129
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-174-077-5

Query Match 31.6%; Score 32.2; DB 4; Length 6129;
 Best Local Similarity 63.6%; Pred. No. 0.059;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | |||| || ||| ||||| || | | | ||| || | ||| | || |||

Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTTTCATTCTG 1604

Qy 65 TTGTCACCTTTCCGAGGA 81
 || ||| ||| | |||

Db 1605 TTCTCAGTTTTTCCTGGA 1621

RESULT 34

PCT-US93-11667-1

; Sequence 1, Application PC/TUS9311667
 ; GENERAL INFORMATION:
 ; APPLICANT: Gregory, R.J., Armentano, D., Couture, L.A., Smith,
 ; APPLICANT: A.E.
 ; TITLE OF INVENTION: GENE THERAPY FOR CYSTIC FIBROSIS
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 STATE STREET, SUITE 510
 ; CITY: BOSTON
 ; STATE: MASSACHUSETTS
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11667
; FILING DATE: 02-DEC-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/985,478
; FILING DATE: 02-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanley, Elizabeth A.
; REGISTRATION NUMBER: 33,505
; REFERENCE/DOCKET NUMBER: NZI-014CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6129 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 133..4572
PCT-US93-11667-1

```

```

Query Match          31.6%; Score 32.2; DB 5; Length 6129;
Best Local Similarity 63.6%; Pred. No. 0.059;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
      || | |||| || ||| ||||| || | | | || | | || | || | ||
Db      1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGGAAGAATTTTCATTCTG 1604

Qy      65 TTGTCACTTTCCGAGGA 81
      || ||| ||| | |||
Db      1605 TTCTCAGTTTTCCTGGA 1621

```

RESULT 35

US-08-466-886-16

```

; Sequence 16, Application US/08466886
; Patent No. 5776677

```

GENERAL INFORMATION:

```

; APPLICANT: Tsui, Lap-Chee
; APPLICANT: Riordan, John R.
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Kerem, Bat-Sheva
; APPLICANT: Collins, Francis S.
; APPLICANT: Iannuzzi, Michael C.
; APPLICANT: Drumm, Mitchell L.
; APPLICANT: Buckwald, Manuel
; TITLE OF INVENTION: Cystic Fibrosis Gene

```

```

; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,886
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldstein, Jorge A.
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 1329.0010006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6130 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 133..4572
US-08-466-886-16

```

```

Query Match          31.6%; Score 32.2; DB 1; Length 6130;
Best Local Similarity 63.6%; Pred. No. 0.059;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
      || | ||| ||   || | |||| ||   | | | ||| || | ||| | || |||
Db      1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTTCATTCTG 1604

Qy      65 TTGTCACTTTCCGAGGA 81
      || ||| ||| | |||
Db      1605 TTCTCAGTTTTCTGGA 1621

```

RESULT 36

US-08-604-488-1

; Sequence 1, Application US/08604488

; Patent No. 5863770

; GENERAL INFORMATION:

; APPLICANT: TSUI, Lap-Chee

; APPLICANT: ROMMENS, Johanna M.

; TITLE OF INVENTION: Stable Propagation of Modified Full

```

; TITLE OF INVENTION: Length Cystic Fibrosis Transmembrane Conductance
Regulator
; TITLE OF INVENTION: Protein cDNA in Heterologous Systems
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell, Seltzer, Park & Gibson
; STREET: 1211 East Morehead Street
; CITY: Charlotte
; STATE: No. 5863770th Carolina
; COUNTRY: U.S.A.
; ZIP: 34009
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/604,488
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/030,081
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Layton, Jr., Samuel G
; REGISTRATION NUMBER: 22,807
; REFERENCE/DOCKET NUMBER: 3477-61
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 704-377-1561
; TELEFAX: 704-334-2014
; TELEX: 57-5102
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6130 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: YES
; ORIGINAL SOURCE:
; TISSUE TYPE: Epithelial
; CELL TYPE: Epithelial cell
; IMMEDIATE SOURCE:
; CLONE: mutant CF gene
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 7
; MAP POSITION: XV2C
; UNITS: bp
US-08-604-488-1

```

```

Query Match          31.6%; Score 32.2; DB 2; Length 6130;
Best Local Similarity 63.6%; Pred. No. 0.059;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy          5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
             || | |||| ||      ||| ||||| ||      | | | ||| || | ||| | || |||

```

```

Db      1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTCATTCTG 1604

Qy      65 TTGTCACTTTCCGAGGA 81
        || ||| ||| | |||
Db      1605 TTCTCAGTTTTCCTGGA 1621

```

RESULT 37

```

US-08-469-461-1
; Sequence 1, Application US/08469461B
; Patent No. 5981178
; GENERAL INFORMATION:
; APPLICANT: Tsui, Lap-Chee
; APPLICANT: Rommins, Johanna M.
; APPLICANT: Kerem, Bat-Sheva
; TITLE OF INVENTION: Introns and Exons of the Cystic Fibrosis Gene and
; TITLE OF INVENTION: Mutations at Various Positions of the Gene
; FILE REFERENCE: 3477-61, 033477/139840
; CURRENT APPLICATION NUMBER: US/08/469,461B
; CURRENT FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 6130
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)..(4572)
US-08-469-461-1

```

```

Query Match          31.6%; Score 32.2; DB 2; Length 6130;
Best Local Similarity 63.6%; Pred. No. 0.059;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
        || | |||| || ||| |||| || | | | ||| || | ||| | || |||
Db      1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTCATTCTG 1604

Qy      65 TTGTCACTTTCCGAGGA 81
        || ||| ||| | |||
Db      1605 TTCTCAGTTTTCCTGGA 1621

```

RESULT 38

```

US-07-890-609-1
; Sequence 1, Application US/07890609C
; Patent No. 6001588
; GENERAL INFORMATION:
; APPLICANT: Tsui, Lap-Chee
; APPLICANT: Rommins, Johanna M.
; APPLICANT: Kerem, Bat-Sheva
; TITLE OF INVENTION: Introns and Exons of the Cystic Fibrosis Gene and
; TITLE OF INVENTION: Mutations at Various Positions of the Gene
; FILE REFERENCE: 3477-61, 033477/139840
; CURRENT APPLICATION NUMBER: US/07/890,609C
; CURRENT FILING DATE: 1992-07-13

```

; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 6130
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)..(4572)
US-07-890-609-1

Query Match 31.6%; Score 32.2; DB 3; Length 6130;
Best Local Similarity 63.6%; Pred. No. 0.059;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
||| ||||| || ||| ||||| || | | | ||| ||| | ||| |||
Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG 1604

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 1605 TTCTCAGTTTTCCTGGA 1621

RESULT 39

US-08-030-081-1

; Sequence 1, Application US/08030081
; Patent No. 6063913

; GENERAL INFORMATION:

; APPLICANT: TSUI, Lap-Chee
; APPLICANT: ROMMENS, Johanna M.
; TITLE OF INVENTION: Stable Propagation of Modified Full
; TITLE OF INVENTION: Length Cystic Fibrosis Transmembrane Conductance
Regulator

; TITLE OF INVENTION: Protein cDNA in Heterologous Systems

; NUMBER OF SEQUENCES: 1

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Bell, Seltzer, Park & Gibson
; STREET: 1211 East Morehead Street
; CITY: Charlotte
; STATE: No. 6063913th Carolina
; COUNTRY: U.S.A.
; ZIP: 34009

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/030,081
; FILING DATE: 19930412
; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Layton, Jr., Samuel G
; REGISTRATION NUMBER: 22,807
; REFERENCE/DOCKET NUMBER: 3477-61

; TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: 704-377-1561
; TELEFAX: 704-334-2014
; TELEX: 57-5102
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6130 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: YES
; ORIGINAL SOURCE:
; TISSUE TYPE: Epithelial
; CELL TYPE: Epithelial cell
; IMMEDIATE SOURCE:
; CLONE: mutant CF gene
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 7
; MAP POSITION: XV2C
; UNITS: bp
US-08-030-081-1

```

```

Query Match          31.6%; Score 32.2; DB 3; Length 6130;
Best Local Similarity 63.6%; Pred. No. 0.059;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
      || | |||| || ||| ||||| || | | ||| || | ||| ||| |||
Db      1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTTCATTCTG 1604

Qy      65 TTGTCACTTTCCGAGGA 81
      || ||| ||| | |||
Db      1605 TTCTCAGTTTTCCTGGA 1621

```

RESULT 40

US-08-469-617-16

```

; Sequence 16, Application US/08469617
; Patent No. 6201107
; GENERAL INFORMATION:
; APPLICANT: Tsui, Lap-Chee
; APPLICANT: Riordan, John R.
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Kerem, Bat-Sheva
; APPLICANT: Collins, Francis S.
; APPLICANT: Iannuzzi, Michael C.
; APPLICANT: Drumm, Mitchell L.
; APPLICANT: Buckwald, Manuel
; TITLE OF INVENTION: Cystic Fibrosis Gene
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA

```

```

; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,617
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldstein, Jorge A.
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 1329.0010008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6130 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 133..4572
US-08-469-617-16

```

```

Query Match          31.6%; Score 32.2; DB 3; Length 6130;
Best Local Similarity 63.6%; Pred. No. 0.059;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
        || | |||| ||      ||| ||||| ||      | | | ||| || | ||| | || |||
Db      1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG 1604

Qy      65 TTGTCACTTTCCGAGGA 81
        || ||| ||| | |||
Db      1605 TTCTCAGTTTCCTGGA 1621

```

RESULT 41

5240846-4

;Patent No. 5240846

```

; APPLICANT: Collins, Francis S.;Wilson, James C.
; TITLE OF INVENTION: GENE THERAPY VECTOR FOR CYSTIC
;FIBROSIS
; NUMBER OF SEQUENCES: 5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/584,275
; FILING DATE: 18-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 399,945
; FILING DATE: 24-AUG-1989
; APPLICATION NUMBER: 401,609
; FILING DATE: 31-AUG-1989

```


;SEQ ID NO:4:
; LENGTH:6146
5240846-4

Query Match 31.6%; Score 32.2; DB 6; Length 6146;
Best Local Similarity 63.6%; Pred. No. 0.059;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| |||| || | | | ||| || | ||| |||
Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 1604

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 1605 TTCTCAGTTTTCTGGA 1621

RESULT 42

US-08-793-618-1

; Sequence 1, Application US/08793618
; Patent No. 6265218

; GENERAL INFORMATION:

; APPLICANT: SEEGER, Stefan
; TITLE OF INVENTION: GENE THERAPY METHOD USING DNA VECTORS
; TITLE OF INVENTION: WITHOUT A SELECTION MARKER GENE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/793,618
; FILING DATE: June 10, 1997
; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/EP95/03027
; FILING DATE: July 31, 1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: DE P 44 28 402.0
; FILING DATE: 11-AUG-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Susan L. Hess
; REGISTRATION NUMBER: 37,350
; REFERENCE/DOCKET NUMBER: BOER 1075 PCT

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 8225 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-793-618-1

Query Match 31.6%; Score 32.2; DB 3; Length 8225;
Best Local Similarity 63.6%; Pred. No. 0.066;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
||| ||||| || ||| ||||| || | | | ||| || | ||| ||| |||
Db 2212 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG 2271

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 2272 TTCTCAGTTTTCTGGA 2288

RESULT 43

US-09-794-431-1

; Sequence 1, Application US/09794431
; Patent No. 6573100
; GENERAL INFORMATION:
; APPLICANT: SEEGER, Stefan
; TITLE OF INVENTION: GENE THERAPY METHOD USING DNA VECTORS
; WITHOUT A SELECTION MARKER GENE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/794,431
; FILING DATE: 27-Feb-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/793,618
; FILING DATE: <Unknown>
; APPLICATION NUMBER: DE P 44 28 402.0
; FILING DATE: 11-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Susan L. Hess
; REGISTRATION NUMBER: 37,350
; REFERENCE/DOCKET NUMBER: BOER 1075 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8225 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-794-431-1

Query Match 31.6%; Score 32.2; DB 4; Length 8225;
Best Local Similarity 63.6%; Pred. No. 0.066;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| |||| || | | | || | | ||| | || |||
Db 2212 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTCATTCTG 2271

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 2272 TTCTCAGTTTTTCCTGGA 2288

RESULT 44

US-08-836-022A-3/c

; Sequence 3, Application US/08836022A

; Patent No. 6001557

; GENERAL INFORMATION:

; APPLICANT: Trustees of the University of Pennsylvania

; APPLICANT: Wilson, James M.

; APPLICANT: Fisher, Krishna J.

; APPLICANT: Chen, Shu-Jen

; APPLICANT: Weitzman, Matthew

; TITLE OF INVENTION: Improved Adenovirus Virus and

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Howson and Howson

; STREET: Spring House Corporate Cntr, P O Box 457

; CITY: Spring House

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 19477

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/836,022A

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/331,381

; FILING DATE: 28-OCT-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Bak, Mary E.

; REGISTRATION NUMBER: 31,215

; REFERENCE/DOCKET NUMBER: GNVPN.008PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9972 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
US-08-836-022A-3

Query Match 31.6%; Score 32.2; DB 3; Length 9972;
Best Local Similarity 63.6%; Pred. No. 0.071;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
||| ||||| || ||| ||||| || | | | ||| || | ||| | ||| |||
Db 7149 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGGAAGAATTTTCATTCTG 7090

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 7089 TTCTCAGTTTTCCTGGA 7073

RESULT 45

US-09-427-048A-3/c

; Sequence 3, Application US/09427048A
; Patent No. 6203975
; GENERAL INFORMATION:
; APPLICANT: Trustees of the University of Pennsylvania
; Wilson, James M.
; Fisher, Krishna J.
; Chen, Shu-Jen
; Weitzman, Matthew
; TITLE OF INVENTION: Improved Adenovirus Virus and
; Methods of Use Thereof
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P O Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/427,048A
; FILING DATE: 21-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/836,022

```

;          FILING DATE: <Unknown>
;  ATTORNEY/AGENT INFORMATION:
;          NAME: Bak, Mary E.
;          REGISTRATION NUMBER: 31,215
;          REFERENCE/DOCKET NUMBER: GNVN.008PCT
;  TELECOMMUNICATION INFORMATION:
;          TELEPHONE: 215-540-9200
;          TELEFAX: 215-540-5818
;  INFORMATION FOR SEQ ID NO: 3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 9972 base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: double
;      TOPOLOGY: unknown
;    MOLECULE TYPE: cDNA
;    SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-427-048A-3

```

```

Query Match          31.6%;  Score 32.2;  DB 3;  Length 9972;
Best Local Similarity 63.6%;  Pred. No. 0.071;
Matches 49;  Conservative 0;  Mismatches 28;  Indels 0;  Gaps 0;

```

```

Qy          5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
             || | |||| ||      ||| ||||| ||  | | | ||| || | ||| | ||| |||
Db          7149 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTCATTCTG 7090

Qy          65 TTGTCACTTTCCGAGGA 81
             || ||| ||| | |||
Db          7089 TTCTCAGTTTTCCTGGA 7073

```

RESULT 46

US-09-423-744A-1

```

; Sequence 1, Application US/09423744A
; Patent No. 6372500
;  GENERAL INFORMATION:
;    APPLICANT: HSC Research and Development Limited Partnership
;    TITLE OF INVENTION: Episomal Expression Cassettes for Gene
;                        Therapy
;    NUMBER OF SEQUENCES: 19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Rockey, Milnamov & Katz, Ltd.
;      STREET: 180 N. Stetson Avenue, Suite 4700
;      CITY: Chicago
;      STATE: Illinois
;      COUNTRY: USA
;      ZIP: 60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/09/423,744A
;      FILING DATE: 12-No. 6372500-1999
;      CLASSIFICATION: <Unknown>
;    PRIOR APPLICATION DATA:

```

```
; APPLICATION NUMBER: PCT/CA98/00478
; FILING DATE: May 14, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Lisa V. Mueller
; REFERENCE/DOCKET NUMBER: DWW6064P0020US
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12143 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Mixture of genomic DNA,
; FEATURE:
; NAME/KEY: enhancer
; LOCATION: 8..2570
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /standard_name= "K18
; Enhancer/Promoter"
; /note= "DNA fragment was obtained by PCR-cloning and minor
; modifications were introduced for the purpose of PCR."
; FEATURE:
; NAME/KEY: intron
; LOCATION: 2571..3318
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /standard_name= "K18 intron 1"
; /note= "DNA fragment was obtained by PCR-cloning and
; modifications were introduced to improve the splicing
; efficiency."
; FEATURE:
; NAME/KEY: enhancer
; LOCATION: 3319..3354
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /standard_name= "Alfalfa mosaic
; virus translational enhancer"
; /note= "Fragment was synthesized chemically."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3355..7948
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /standard_name= "CFTR cDNA"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7949..7984
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /standard_name= "pBluescript II
; KS(+) multiple cloning site"
; FEATURE:
; NAME/KEY: intron
; LOCATION: 8507..8572
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /standard_name= "SV40 small t
; antigen intron"
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 9178..9212
; IDENTIFICATION METHOD:
```

```

;         OTHER INFORMATION: /standard_name= "SV40
;         polyadenylation signal"
;     FEATURE:
;         NAME/KEY:  polyA_signal
;         LOCATION: 12021..12055
;         IDENTIFICATION METHOD:
;         OTHER INFORMATION: /standard_name= "SV40
;         polyadenylation signal"
;     FEATURE:
;         NAME/KEY:  rep_origin
;         LOCATION: 9562..10205
;         IDENTIFICATION METHOD:
;         OTHER INFORMATION: /standard_name= "pUC origin of
;         replication"
;     FEATURE:
;         NAME/KEY:  misc_feature
;         LOCATION: 11283..11353
;         IDENTIFICATION METHOD:
;         OTHER INFORMATION: /standard_name= "Ampicillin
;         resistance gene"
;     FEATURE:
;         NAME/KEY:  misc_feature
;         LOCATION: 11345..11800
;         IDENTIFICATION METHOD:
;         OTHER INFORMATION: /standard_name= "f1 single strand
;         DNA origin"
;     SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-423-744A-1

```

```

Query Match          31.6%;  Score 32.2;  DB 4;  Length 12143;
Best Local Similarity 63.6%;  Pred. No. 0.077;
Matches 49;  Conservative 0;  Mismatches 28;  Indels 0;  Gaps 0;

```

```

Qy          5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
             || | |||| ||      ||| ||||| ||  | | | ||| || | ||| | || |||
Db          4767 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGGAAGAATTCATTCTG 4826

Qy          65 TTGTCACTTTCCGAGGA 81
             || ||| ||| | |||
Db          4827 TTCTCAGTTTTCCTGGA 4843

```

RESULT 47

US-08-469-461-3

```

; Sequence 3, Application US/08469461B
; Patent No. 5981178
; GENERAL INFORMATION:
; APPLICANT: Tsui, Lap-Chee
; APPLICANT: Rommins, Johanna M.
; APPLICANT: Kerem, Bat-Sheva
; TITLE OF INVENTION: Introns and Exons of the Cystic Fibrosis Gene and
; TITLE OF INVENTION: Mutations at Various Positions of the Gene
; FILE REFERENCE: 3477-61, 033477/139840
; CURRENT APPLICATION NUMBER: US/08/469,461B
; CURRENT FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0

```

; SEQ ID NO 3
; LENGTH: 22846
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-469-461-3

Query Match 31.6%; Score 32.2; DB 2; Length 22846;
Best Local Similarity 63.6%; Pred. No. 0.099;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
||| ||||| || ||| ||||| || ||| ||||| || ||| ||| |||
Db 8874 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTCACTCTG 8933

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 8934 TTCTCAGTTTTCTGGA 8950

RESULT 48

US-07-890-609-3

; Sequence 3, Application US/07890609C
; Patent No. 6001588
; GENERAL INFORMATION:
; APPLICANT: Tsui, Lap-Chee
; APPLICANT: Rommins, Johanna M.
; APPLICANT: Kerem, Bat-Sheva
; TITLE OF INVENTION: Introns and Exons of the Cystic Fibrosis Gene and
; TITLE OF INVENTION: Mutations at Various Positions of the Gene
; FILE REFERENCE: 3477-61, 033477/139840
; CURRENT APPLICATION NUMBER: US/07/890,609C
; CURRENT FILING DATE: 1992-07-13
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 22846
; TYPE: DNA
; ORGANISM: Homo sapiens
US-07-890-609-3

Query Match 31.6%; Score 32.2; DB 3; Length 22846;
Best Local Similarity 63.6%; Pred. No. 0.099;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
||| ||||| || ||| ||||| || ||| ||||| || ||| ||| |||
Db 8874 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTCACTCTG 8933

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 8934 TTCTCAGTTTTCTGGA 8950

RESULT 49

US-09-252-991A-1019/c

; Sequence 1019, Application US/09252991A
; Patent No. 6551795


```
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1019
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1019
```

```
Query Match          29.6%; Score 30.2; DB 4; Length 1251;
Best Local Similarity 56.6%; Pred. No. 0.17;
Matches 56; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
```

```
Qy      3 GGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGAC 62
        ||||| | | ||| ||| || || | | || | |
Db      1067 GGTAGGTGTTGCCCTTCACCACCAGGTCGTCGGCCTCGTAGCAATAGAAGCCGTACCAGC 1008

Qy      63 TGTGTCACCTTCCGAGGAGAAACAAGCTGTCCTGGAGGC 101
        ||| | | | ||| || || | | || | |||
Db      1007 TGTGACGATGGTCGAGTCGATCAGCCAGCCCTTGGGGC 969
```

RESULT 50

```
US-09-252-991A-1036
; Sequence 1036, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1036
; LENGTH: 2847
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1036
```

```
Query Match          29.6%; Score 30.2; DB 4; Length 2847;
Best Local Similarity 56.6%; Pred. No. 0.23;
```

Matches 56; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```
Qy      3 GGTAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAGAC 62
        ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1818 GGTAGGTGTTGCCCTTCACCACCAGGTCGTCGGCCTCGTAGCAATAGAAGCCGTACCAGC 1877

Qy      63 TGTGTGTCACCTTCCGAGGAGACAAGCTGTCCTGGAGGC 101
        ||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1878 TGTCGACGATGGTCGAGTCGATCACCCAGCCCTTGGGGC 1916
```

Search completed: April 29, 2004, 17:08:31

Job time : 11.3337 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 29, 2004, 17:06:46 ; Search time 99.1938 Seconds
(without alignments)
4651.434 Million cell updates/sec

Title: US-09-989-981A-9_COPY_3_104
Perfect score: 102
Sequence: 1 ctggtaggtgagatctctga.....aacaagctgtcctggaggcc 102

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2936184 seqs, 2261732022 residues

Total number of hits satisfying chosen parameters: 5872368

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : Published Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result 8
Query

	No.	Score	Match	Length	DB	ID	Description
c	1	102	100.0	2019	10	US-09-989-981A-3	Sequence 3, Appli
	2	102	100.0	6043	10	US-09-989-981A-9	Sequence 9, Appli
c	3	87.6	85.9	2669	10	US-09-989-981A-7	Sequence 7, Appli
	4	32.2	31.6	180	9	US-09-864-761-27920	Sequence 27920, A
	5	32.2	31.6	240	12	US-10-441-643-1	Sequence 1, Appli
	6	32.2	31.6	420	9	US-09-756-095-64	Sequence 64, Appl
	7	32.2	31.6	420	10	US-09-941-492-64	Sequence 64, Appl
	8	32.2	31.6	420	10	US-09-756-096A-64	Sequence 64, Appl
	9	32.2	31.6	420	10	US-09-838-858-64	Sequence 64, Appl
	10	32.2	31.6	494	9	US-09-864-761-11433	Sequence 11433, A
	11	32.2	31.6	831	16	US-10-300-683-247	Sequence 247, App
	12	32.2	31.6	3069	9	US-09-756-095-105	Sequence 105, App
	13	32.2	31.6	3069	10	US-09-941-492-105	Sequence 105, App
	14	32.2	31.6	3069	10	US-09-756-096A-105	Sequence 105, App
	15	32.2	31.6	3069	10	US-09-838-858-105	Sequence 105, App
	16	32.2	31.6	4191	16	US-10-367-507-1	Sequence 1, Appli
	17	32.2	31.6	4311	16	US-10-367-507-8	Sequence 8, Appli
	18	32.2	31.6	4347	16	US-10-367-507-6	Sequence 6, Appli
	19	32.2	31.6	4347	16	US-10-367-507-7	Sequence 7, Appli
	20	32.2	31.6	4368	16	US-10-367-507-5	Sequence 5, Appli
	21	32.2	31.6	4371	16	US-10-367-507-4	Sequence 4, Appli
	22	32.2	31.6	4410	16	US-10-367-507-3	Sequence 3, Appli
	23	32.2	31.6	4419	16	US-10-367-507-2	Sequence 2, Appli
	24	32.2	31.6	5635	15	US-10-161-539-3	Sequence 3, Appli
	25	32.2	31.6	6126	10	US-09-982-315-3	Sequence 3, Appli
	26	32.2	31.6	6129	9	US-09-782-378A-24	Sequence 24, Appl
	27	32.2	31.6	6129	10	US-09-982-315-1	Sequence 1, Appli
	28	32.2	31.6	6129	10	US-09-982-315-5	Sequence 5, Appli
	29	32.2	31.6	6129	15	US-10-161-539-1	Sequence 1, Appli
c	30	29	28.4	2310	16	US-10-369-493-37753	Sequence 37753, A
	31	28.4	27.8	799	13	US-10-027-632-172129	Sequence 172129,
	32	28.4	27.8	799	16	US-10-027-632-172129	Sequence 172129,
c	33	28.2	27.6	548	9	US-09-925-299-368	Sequence 368, App
c	34	28.2	27.6	548	10	US-09-925-299-368	Sequence 368, App
c	35	28	27.5	432	9	US-09-934-814-4	Sequence 4, Appli
c	36	28	27.5	432	15	US-10-142-465-4	Sequence 4, Appli
c	37	28	27.5	525	9	US-09-934-814-1	Sequence 1, Appli
c	38	28	27.5	525	15	US-10-142-465-1	Sequence 1, Appli
c	39	28	27.5	540	9	US-09-934-814-7	Sequence 7, Appli
c	40	28	27.5	540	15	US-10-142-465-7	Sequence 7, Appli
c	41	28	27.5	795	9	US-09-934-814-10	Sequence 10, Appl
c	42	28	27.5	795	15	US-10-142-465-10	Sequence 10, Appl
c	43	28	27.5	1630	13	US-10-399-456-3	Sequence 3, Appli
	44	28	27.5	48244	12	US-10-052-482-166	Sequence 166, App
	45	27.8	27.3	987	13	US-10-027-632-265948	Sequence 265948,
	46	27.8	27.3	987	13	US-10-027-632-265949	Sequence 265949,
	47	27.8	27.3	987	13	US-10-027-632-265950	Sequence 265950,
	48	27.8	27.3	987	13	US-10-027-632-265951	Sequence 265951,
	49	27.8	27.3	987	16	US-10-027-632-265948	Sequence 265948,
	50	27.8	27.3	987	16	US-10-027-632-265949	Sequence 265949,

ALIGNMENTS

RESULT 1

US-09-989-981A-3/c

```
; Sequence 3, Application US/09989981A
; Publication No. US20030049730A1
; GENERAL INFORMATION:
; APPLICANT: Hobbs, Helen H.
; APPLICANT: Shan, Bei
; APPLICANT: Barnes, Robert
; APPLICANT: Tian, Hui
; APPLICANT: Tularik Inc.
; APPLICANT: Board of Regents, The University of Texas System
; TITLE OF INVENTION: ABCG5 and ABCG8: Compositions and Methods of Use
; FILE REFERENCE: 018781-007320US
; CURRENT APPLICATION NUMBER: US/09/989,981A
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 60/252,235
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/253,645
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2019
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2019)
; OTHER INFORMATION: mouse ABCG8 (mABCG8)
```

US-09-989-981A-3

```
Query Match          100.0%; Score 102; DB 10; Length 2019;
Best Local Similarity 100.0%; Pred. No. 8e-27;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      165 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAG 106

Qy      61 ACTGTTGTCACTTTCGAGGAGAACAAGCTGTCCTGGAGGCC 102
          ||||||||||||||||||||||||||||
Db      105 ACTGTTGTCACTTTCGAGGAGAACAAGCTGTCCTGGAGGCC 64
```

RESULT 2

US-09-989-981A-9

```
; Sequence 9, Application US/09989981A
; Publication No. US20030049730A1
; GENERAL INFORMATION:
; APPLICANT: Hobbs, Helen H.
; APPLICANT: Shan, Bei
; APPLICANT: Barnes, Robert
; APPLICANT: Tian, Hui
; APPLICANT: Tularik Inc.
; APPLICANT: Board of Regents, The University of Texas System
; TITLE OF INVENTION: ABCG5 and ABCG8: Compositions and Methods of Use
; FILE REFERENCE: 018781-007320US
```

```
; CURRENT APPLICATION NUMBER: US/09/989,981A
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 60/252,235
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/253,645
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 6043
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: ABCG8 exon 2 (reverse strand) through ABCG5 exon 2
; OTHER INFORMATION: (forward strand)
US-09-989-981A-9
```

```
Query Match          100.0%; Score 102; DB 10; Length 6043;
Best Local Similarity 100.0%; Pred. No. 1.1e-26;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 62

Qy      61 ACTGTTGTCACTTTCGAGGAGAACAAGCTGTCCTGGAGGCC 102
          ||||||||||||||||||||||||||||
Db      63 ACTGTTGTCACTTTCGAGGAGAACAAGCTGTCCTGGAGGCC 104
```

RESULT 3

```
US-09-989-981A-7/c
; Sequence 7, Application US/09989981A
; Publication No. US20030049730A1
; GENERAL INFORMATION:
; APPLICANT: Hobbs, Helen H.
; APPLICANT: Shan, Bei
; APPLICANT: Barnes, Robert
; APPLICANT: Tian, Hui
; APPLICANT: Tularik Inc.
; APPLICANT: Board of Regents, The University of Texas System
; TITLE OF INVENTION: ABCG5 and ABCG8: Compositions and Methods of Use
; FILE REFERENCE: 018781-007320US
; CURRENT APPLICATION NUMBER: US/09/989,981A
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 60/252,235
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/253,645
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2669
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
```

; LOCATION: (100)..(2121)
; OTHER INFORMATION: human ABCG8 (hABCG8)
US-09-989-981A-7

Query Match 85.9%; Score 87.6; DB 10; Length 2669;
Best Local Similarity 91.2%; Pred. No. 1.5e-21;
Matches 93; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
||||| ||| |||||||||||| ||||| || ||||||||||||||||
Db 264 CTGGTAGTTGAGGTCTCTGACCTCCAGGGTGTGGGCTGGCCACTGTAGGTGAAGTACAG 205

Qy 61 ACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGGAGGCC 102
|||||||||| |||||||||||| || |||||||||
Db 204 GCTGTTGTCACTTTTCAGAGGAGAACAATCTATCCTGGAGGCC 163

RESULT 4

US-09-864-761-27920

; Sequence 27920, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES
USEFUL FOR

; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

; FILE REFERENCE: Aeomica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 27920
; LENGTH: 180
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000111.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: NT HIT: gill1422155, EVALUE 4.00e-97
; OTHER INFORMATION: SWISSPROT HIT: P13569, EVALUE 6.00e-30
; OTHER INFORMATION: EST_HUMAN HIT: AA524439.1, EVALUE 8.00e-59
US-09-864-761-27920

```

```

Query Match          31.6%; Score 32.2; DB 9; Length 180;
Best Local Similarity 63.6%; Pred. No. 0.11;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
        || | |||| ||      ||| ||||| ||      | | | ||| || | ||| | ||| |||
Db      9 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTTCATTCTG 68

Qy      65 TTGTCACTTTCCGAGGA 81
        || ||| ||| | |||
Db      69 TTCTCAGTTTTCCTGGA 85

```

RESULT 5

US-10-441-643-1

```

; Sequence 1, Application US/10441643
; Publication No. US20040072208A1
; GENERAL INFORMATION:
; APPLICANT: Warthoe, Peter
; TITLE OF INVENTION: Surface Acoustic Wave Sensors and Method for Detecting
Target
; TITLE OF INVENTION: Analytes
; FILE REFERENCE: A-71523
; CURRENT APPLICATION NUMBER: US/10/441,643
; CURRENT FILING DATE: 2003-05-20
; PRIOR APPLICATION NUMBER: US 60/383,247
; PRIOR FILING DATE: 2002-05-23

```


; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 240
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-441-643-1

Query Match 31.6%; Score 32.2; DB 12; Length 240;
Best Local Similarity 63.6%; Pred. No. 0.12;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
||| ||||| || ||||| || ||| ||| ||| ||| ||| |||
Db 51 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 110

Qy 65 TTGTCACTTTCCGAGGA 81
||| ||| ||| |||
Db 111 TTCTCAGTTTTCCTGGA 127

RESULT 6

US-09-756-095-64

; Sequence 64, Application US/09756095
; Patent No. US20020115207A1
; GENERAL INFORMATION:
; APPLICANT: Mitchell, Lloyd G.
; APPLICANT: Garcia-Blanco, Mariano A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR USE IN
; TITLE OF INVENTION: SPLICEOSOME MEDIATED RNA TRANS-SPLICING
; FILE REFERENCE: A31304-B-A 072874.0134
; CURRENT APPLICATION NUMBER: US/09/756,095
; CURRENT FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 09/158,863
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 09/133,717
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 09/087,233
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 08/766,354
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: 60/008,317
; PRIOR FILING DATE: 1995-12-07
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: trans spliced product comprising cystic fibrosis
; OTHER INFORMATION: transmembrane regulator-derived sequences and His
; OTHER INFORMATION: tag sequences
US-09-756-095-64

Query Match 31.6%; Score 32.2; DB 9; Length 420;
Best Local Similarity 63.6%; Pred. No. 0.14;

Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
      || | |||| ||   ||| ||||| ||   | | | ||| || | ||| | || |||
Db     128 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 187

Qy      65 TTGTCACTTTCCGAGGA 81
      || ||| ||| | |||
Db     188 TTCTCAGTTTTCCTGGA 204
```

RESULT 7

US-09-941-492-64

; Sequence 64, Application US/09941492

; Publication No. US20030027250A1

; GENERAL INFORMATION:

; APPLICANT: Mitchell, Lloyd

; APPLICANT: Garcia-Blanco, Mariano M.

; APPLICANT: Puttaraju, Madaiah

; APPLICANT: Mansfield, Gary S.

; TITLE OF INVENTION: METHODS OF COMPOSITIONS FOR USE IN

; TITLE OF INVENTION: SPLICEOSOME MEDIATED RNA TRANS-SPLICING

; FILE REFERENCE: A31304-BAE (072874.0156)

; CURRENT APPLICATION NUMBER: US/09/941,492

; CURRENT FILING DATE: 2002-04-01

; PRIOR APPLICATION NUMBER: 09/838,858

; PRIOR FILING DATE: 2001-04-20

; PRIOR APPLICATION NUMBER: 09/756,096

; PRIOR FILING DATE: 2001-01-08

; PRIOR APPLICATION NUMBER: 09/158,863

; PRIOR FILING DATE: 1998-09-23

; PRIOR APPLICATION NUMBER: 09/133,717

; PRIOR FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: 09/087,233

; PRIOR FILING DATE: 1998-05-28

; PRIOR APPLICATION NUMBER: 08/766,354

; PRIOR FILING DATE: 1996-12-13

; NUMBER OF SEQ ID NOS: 125

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 64

; LENGTH: 420

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Trans-spliced product comprising cystic fibrosis

; OTHER INFORMATION: transmembrane regulator-derived sequences and His

; OTHER INFORMATION: tag sequences

US-09-941-492-64

Query Match 31.6%; Score 32.2; DB 10; Length 420;

Best Local Similarity 63.6%; Pred. No. 0.14;

Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
      || | |||| ||   ||| ||||| ||   | | | ||| || | ||| | || |||
Db     128 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 187
```

Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 188 TTCTCAGTTTTCCTGGA 204

RESULT 8

US-09-756-096A-64

; Sequence 64, Application US/09756096A
 ; Publication No. US20030077754A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mitchell, Lloyd G.
 ; APPLICANT: Garcia-Blanco, Mariano A.
 ; APPLICANT: Puttaraju, Madaiah
 ; APPLICANT: Mansfield, Gary S.
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR USE IN
 ; TITLE OF INVENTION: SPLICEOSOME MEDIATED RNA TRANS-SPLICING
 ; FILE REFERENCE: A31304-B-A-B 072874.0135
 ; CURRENT APPLICATION NUMBER: US/09/756,096A
 ; CURRENT FILING DATE: 2001-01-08
 ; PRIOR APPLICATION NUMBER: 09/158,863
 ; PRIOR FILING DATE: 1998-09-23
 ; PRIOR APPLICATION NUMBER: 09/133,717
 ; PRIOR FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: 09/087,233
 ; PRIOR FILING DATE: 1998-05-28
 ; PRIOR APPLICATION NUMBER: 08/766,354
 ; PRIOR FILING DATE: 1996-12-13
 ; PRIOR APPLICATION NUMBER: 60/008,317
 ; PRIOR FILING DATE: 1995-12-15
 ; NUMBER OF SEQ ID NOS: 105
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 64
 ; LENGTH: 420
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: trans-spliced product comprising cystic fibrosis
 ; OTHER INFORMATION: transmembrane regulator-derived sequences and His
 ; OTHER INFORMATION: tag sequence
 US-09-756-096A-64

Query Match 31.6%; Score 32.2; DB 10; Length 420;
 Best Local Similarity 63.6%; Pred. No. 0.14;
 Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
 || | ||| || ||| |||| || | | | ||| || | ||| | || |||
 Db 128 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTCATTCTG 187

 Qy 65 TTGTCACTTTCCGAGGA 81
 || ||| ||| | |||
 Db 188 TTCTCAGTTTTCCTGGA 204

RESULT 9

US-09-838-858-64

; Sequence 64, Application US/09838858

```
; Publication No. US20030148937A1
; GENERAL INFORMATION:
; APPLICANT: Mansfield, Gary S.
; APPLICANT: Mitchell, Lloyd G.
; APPLICANT: Garcia-Blanco, Mariano A.
; APPLICANT: Walsh, Christopher E.
; APPLICANT: Chao, Hengjun
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR USE IN
; TITLE OF INVENTION: SPLICEOSOME MEDIATED RNA TRANS-SPLICING
; FILE REFERENCE: A31304-BAD 072874.01
; CURRENT APPLICATION NUMBER: US/09/838,858
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 09/756,096
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 09/158,863
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 09/133,717
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 09/087,233
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 08/766,354
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: 60/008,317
; PRIOR FILING DATE: 1995-12-15
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Trans-spliced product containing cystic fibrosis
; OTHER INFORMATION: transmembrane regulator-derived sequences and
; OTHER INFORMATION: His-tag sequence
US-09-838-858-64
```

```
Query Match          31.6%; Score 32.2; DB 10; Length 420;
Best Local Similarity 63.6%; Pred. No. 0.14;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
```

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
      || | |||| ||      ||| ||||| ||      | | | ||| || | ||| | || |||
Db     128 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 187

Qy      65 TTGTCACTTTCCGAGGA 81
      || ||| ||| | |||
Db     188 TTCTCAGTTTTCCTGGA 204
```

```
RESULT 10
US-09-864-761-11433
; Sequence 11433, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
```

; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES
USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 11433
; LENGTH: 494
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000111.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1

US-09-864-761-11433

Query Match 31.6%; Score 32.2; DB 9; Length 494;
Best Local Similarity 63.6%; Pred. No. 0.14;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
          || | |||| ||   ||| ||||| ||   | | | ||| || | ||| | || |||
Db      280 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTTTCATTCTG 339

Qy      65 TTGTCACTTTCCGAGGA 81
          || ||| ||| | |||
Db      340 TTCTCAGTTTTTCCTGGA 356
```

RESULT 11

US-10-300-683-247

; Sequence 247, Application US/10300683
; Publication No. US20030235834A1
; GENERAL INFORMATION:
; APPLICANT: Dunlop, Charles L.M.
; APPLICANT: Weisel, James M.
; TITLE OF INVENTION: APPROACHES TO IDENTIFY CYSTIC FIBROSIS
; FILE REFERENCE: CHARDUN.010A
; CURRENT APPLICATION NUMBER: US/10/300,683
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 60/333,531
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 554
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 247
; LENGTH: 831
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Diagnostic Oligonucleotide

US-10-300-683-247

Query Match 31.6%; Score 32.2; DB 16; Length 831;
Best Local Similarity 63.6%; Pred. No. 0.16;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
          || | |||| ||   ||| ||||| ||   | | | ||| || | ||| | || |||
Db      328 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTTTCATTCTG 387

Qy      65 TTGTCACTTTCCGAGGA 81
          || ||| ||| | |||
Db      388 TTCTCAGTTTTTCCTGGA 404
```

RESULT 12

US-09-756-095-105

; Sequence 105, Application US/09756095
; Patent No. US20020115207A1
; GENERAL INFORMATION:
; APPLICANT: Mitchell, Lloyd G.

```

; APPLICANT: Garcia-Blanco, Mariano A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR USE IN
; TITLE OF INVENTION: SPLICEOSOME MEDIATED RNA TRANS-SPLICING
; FILE REFERENCE: A31304-B-A 072874.0134
; CURRENT APPLICATION NUMBER: US/09/756,095
; CURRENT FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 09/158,863
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 09/133,717
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 09/087,233
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 08/766,354
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: 60/008,317
; PRIOR FILING DATE: 1995-12-07
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 3069
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CFTR PTM sequence
US-09-756-095-105

```

```

Query Match          31.6%; Score 32.2; DB 9; Length 3069;
Best Local Similarity 63.6%; Pred. No. 0.23;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
      || | |||| ||      ||| ||||| ||      | | | ||| || | ||| | || |||
Db      21 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG 80

Qy      65 TTGTCACTTTCCGAGGA 81
      || ||| ||| | |||
Db      81 TTCTCAGTTTTCCTGGA 97

```

RESULT 13

```

US-09-941-492-105
; Sequence 105, Application US/09941492
; Publication No. US20030027250A1
; GENERAL INFORMATION:
; APPLICANT: Mitchell, Lloyd
; APPLICANT: Garcia-Blanco, Mariano M.
; APPLICANT: Puttaraju, Madaiah
; APPLICANT: Mansfield, Gary S.
; TITLE OF INVENTION: METHODS OF COMPOSITIONS FOR USE IN
; TITLE OF INVENTION: SPLICEOSOME MEDIATED RNA TRANS-SPLICING
; FILE REFERENCE: A31304-BAE (072874.0156)
; CURRENT APPLICATION NUMBER: US/09/941,492
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/838,858
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 09/756,096
; PRIOR FILING DATE: 2001-01-08

```

```
; PRIOR APPLICATION NUMBER: 09/158,863
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 09/133,717
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 09/087,233
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 08/766,354
; PRIOR FILING DATE: 1996-12-13
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 3069
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CFTR PTM sequence
US-09-941-492-105
```

```
Query Match          31.6%; Score 32.2; DB 10; Length 3069;
Best Local Similarity 63.6%; Pred. No. 0.23;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
```

```
Qy          5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
              ||| ||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| |||
Db          21 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTTCATTCTG 80

Qy          65 TTGTCACTTTCCGAGGA 81
              || ||| ||| | |||
Db          81 TTCTCAGTTTTCCTGGA 97
```

RESULT 14

US-09-756-096A-105

```
; Sequence 105, Application US/09756096A
; Publication No. US20030077754A1
; GENERAL INFORMATION:
; APPLICANT: Mitchell, Lloyd G.
; APPLICANT: Garcia-Blanco, Mariano A.
; APPLICANT: Puttaraju, Madaiah
; APPLICANT: Mansfield, Gary S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR USE IN
; TITLE OF INVENTION: SPLICEOSOME MEDIATED RNA TRANS-SPLICING
; FILE REFERENCE: A31304-B-A-B 072874.0135
; CURRENT APPLICATION NUMBER: US/09/756,096A
; CURRENT FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 09/158,863
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 09/133,717
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 09/087,233
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 08/766,354
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: 60/008,317
; PRIOR FILING DATE: 1995-12-15
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSEQ for Windows Version 4.0
```


; SEQ ID NO 105
; LENGTH: 3069
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CFTR PTM sequence
US-09-756-096A-105

Query Match 31.6%; Score 32.2; DB 10; Length 3069;
Best Local Similarity 63.6%; Pred. No. 0.23;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| |||| || | | ||| || | ||| || |||
Db 21 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 80

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 81 TTCTCAGTTTTCCTGGA 97

RESULT 15

US-09-838-858-105
; Sequence 105, Application US/09838858
; Publication No. US20030148937A1
; GENERAL INFORMATION:
; APPLICANT: Mansfield, Gary S.
; APPLICANT: Mitchell, Lloyd G.
; APPLICANT: Garcia-Blanco, Mariano A.
; APPLICANT: Walsh, Christopher E.
; APPLICANT: Chao, Hengjun
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR USE IN
; TITLE OF INVENTION: SPLICEOSOME MEDIATED RNA TRANS-SPlicing
; FILE REFERENCE: A31304-BAD 072874.01
; CURRENT APPLICATION NUMBER: US/09/838,858
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 09/756,096
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 09/158,863
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 09/133,717
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 09/087,233
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 08/766,354
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: 60/008,317
; PRIOR FILING DATE: 1995-12-15
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 3069
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CFTR PTM sequence
US-09-838-858-105

Query Match 31.6%; Score 32.2; DB 10; Length 3069;
Best Local Similarity 63.6%; Pred. No. 0.23;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| |||| || | | | ||| || | ||| ||| |||

Db 21 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 80

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||

Db 81 TTCTCAGTTTTCTCTGGA 97

RESULT 16

US-10-367-507-1

; Sequence 1, Application US/10367507
; Publication No. US20030235885A1
; GENERAL INFORMATION:
; APPLICANT: Welsh, Michael J.
; APPLICANT: Ostedgaard, Lynda S.
; APPLICANT: Zabner, Joseph
; TITLE OF INVENTION: CFTR WITH A PARTIALLY DELETED R DOMAIN
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AP35027 (072419.0117)
; CURRENT APPLICATION NUMBER: US/10/367,507
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: 60/358,074
; PRIOR FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4191
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)...(4191)
US-10-367-507-1

Query Match 31.6%; Score 32.2; DB 16; Length 4191;
Best Local Similarity 63.6%; Pred. No. 0.25;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
|| | |||| || ||| |||| || | | | ||| || | ||| ||| |||

Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 1604

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||

Db 1605 TTCTCAGTTTTCTCTGGA 1621

RESULT 17

US-10-367-507-8

; Sequence 8, Application US/10367507
; Publication No. US20030235885A1

```
; GENERAL INFORMATION:
; APPLICANT: Welsh, Michael J.
; APPLICANT: Ostedgaard, Lynda S.
; APPLICANT: Zabner, Joseph
; TITLE OF INVENTION: CFTR WITH A PARTIALLY DELETED R DOMAIN
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AP35027 (072419.0117)
; CURRENT APPLICATION NUMBER: US/10/367,507
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: 60/358,074
; PRIOR FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 4311
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)...(4311)
US-10-367-507-8
```

```
Query Match          31.6%; Score 32.2; DB 16; Length 4311;
Best Local Similarity 63.6%; Pred. No. 0.25;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
```

```
Qy          5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
             || | |||| ||   ||| |||| ||   | | | ||| || | ||| | || |||
Db          1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTTCATTCTG 1604

Qy          65 TTGTCACTTTCCGAGGA 81
             || ||| ||| | |||
Db          1605 TTCTCAGTTTTTCCTGGA 1621
```

RESULT 18

US-10-367-507-6

```
; Sequence 6, Application US/10367507
; Publication No. US20030235885A1
; GENERAL INFORMATION:
; APPLICANT: Welsh, Michael J.
; APPLICANT: Ostedgaard, Lynda S.
; APPLICANT: Zabner, Joseph
; TITLE OF INVENTION: CFTR WITH A PARTIALLY DELETED R DOMAIN
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AP35027 (072419.0117)
; CURRENT APPLICATION NUMBER: US/10/367,507
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: 60/358,074
; PRIOR FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 4347
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
```

; NAME/KEY: CDS
; LOCATION: (133)...(4347)
US-10-367-507-6

Query Match 31.6%; Score 32.2; DB 16; Length 4347;
Best Local Similarity 63.6%; Pred. No. 0.25;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
          || | |||| ||   ||| ||||| ||   | | | ||| || |   ||| | ||   |||
Db      1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTCATTCTG 1604

Qy      65 TTGTCACTTTCCGAGGA 81
          || ||| ||| |   |||
Db      1605 TTCTCAGTTTTTCCTGGA 1621
```

RESULT 19

US-10-367-507-7
; Sequence 7, Application US/10367507
; Publication No. US20030235885A1
; GENERAL INFORMATION:
; APPLICANT: Welsh, Michael J.
; APPLICANT: Ostedgaard, Lynda S.
; APPLICANT: Zabner, Joseph
; TITLE OF INVENTION: CFTR WITH A PARTIALLY DELETED R DOMAIN
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AP35027 (072419.0117)
; CURRENT APPLICATION NUMBER: US/10/367,507
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: 60/358,074
; PRIOR FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 4347
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)...(4347)
US-10-367-507-7

Query Match 31.6%; Score 32.2; DB 16; Length 4347;
Best Local Similarity 63.6%; Pred. No. 0.25;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
          || | |||| ||   ||| ||||| ||   | | | ||| || |   ||| | ||   |||
Db      1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTCATTCTG 1604

Qy      65 TTGTCACTTTCCGAGGA 81
          || ||| ||| |   |||
Db      1605 TTCTCAGTTTTTCCTGGA 1621
```

RESULT 20

US-10-367-507-5

```
; Sequence 5, Application US/10367507
; Publication No. US20030235885A1
; GENERAL INFORMATION:
; APPLICANT: Welsh, Michael J.
; APPLICANT: Ostedgaard, Lynda S.
; APPLICANT: Zabner, Joseph
; TITLE OF INVENTION: CFTR WITH A PARTIALLY DELETED R DOMAIN
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AP35027 (072419.0117)
; CURRENT APPLICATION NUMBER: US/10/367,507
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: 60/358,074
; PRIOR FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4368
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)...(4368)
US-10-367-507-5
```

```
Query Match          31.6%; Score 32.2; DB 16; Length 4368;
Best Local Similarity 63.6%; Pred. No. 0.25;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
```

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
      || | ||| || || | ||| || | | | || | | || | || | ||
Db      1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTATTCTG 1604

Qy      65 TTGTCACTTTCCGAGGA 81
      || ||| ||| | |||
Db      1605 TTCTCAGTTTTTCCTGGA 1621
```

RESULT 21

US-10-367-507-4

```
; Sequence 4, Application US/10367507
; Publication No. US20030235885A1
; GENERAL INFORMATION:
; APPLICANT: Welsh, Michael J.
; APPLICANT: Ostedgaard, Lynda S.
; APPLICANT: Zabner, Joseph
; TITLE OF INVENTION: CFTR WITH A PARTIALLY DELETED R DOMAIN
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AP35027 (072419.0117)
; CURRENT APPLICATION NUMBER: US/10/367,507
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: 60/358,074
; PRIOR FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 4371
```

; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)...(4371)
US-10-367-507-4

Query Match 31.6%; Score 32.2; DB 16; Length 4371;
Best Local Similarity 63.6%; Pred. No. 0.25;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
||| ||||| || ||||| || ||| ||||| ||| ||| |||
Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 1604

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 1605 TTCTCAGTTTCCTGGA 1621

RESULT 22

US-10-367-507-3

; Sequence 3, Application US/10367507
; Publication No. US20030235885A1
; GENERAL INFORMATION:
; APPLICANT: Welsh, Michael J.
; APPLICANT: Ostedgaard, Lynda S.
; APPLICANT: Zabner, Joseph
; TITLE OF INVENTION: CFTR WITH A PARTIALLY DELETED R DOMAIN
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AP35027 (072419.0117)
; CURRENT APPLICATION NUMBER: US/10/367,507
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: 60/358,074
; PRIOR FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 4410
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)...(4410)
US-10-367-507-3

Query Match 31.6%; Score 32.2; DB 16; Length 4410;
Best Local Similarity 63.6%; Pred. No. 0.25;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
||| ||||| || ||||| || ||| ||||| ||| ||| |||
Db 1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTCATTCTG 1604

Qy 65 TTGTCACTTTCCGAGGA 81
|| ||| ||| | |||
Db 1605 TTCTCAGTTTCCTGGA 1621

RESULT 23

US-10-367-507-2

```
; Sequence 2, Application US/10367507
; Publication No. US20030235885A1
; GENERAL INFORMATION:
; APPLICANT: Welsh, Michael J.
; APPLICANT: Ostedgaard, Lynda S.
; APPLICANT: Zabner, Joseph
; TITLE OF INVENTION: CFTR WITH A PARTIALLY DELETED R DOMAIN
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AP35027 (072419.0117)
; CURRENT APPLICATION NUMBER: US/10/367,507
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: 60/358,074
; PRIOR FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 4419
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)...(4419)
```

US-10-367-507-2

```
Query Match          31.6%; Score 32.2; DB 16; Length 4419;
Best Local Similarity 63.6%; Pred. No. 0.25;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
```

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
      || | ||| ||   || | ||| ||   | | | ||| || | ||| | || |||
Db      1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTCATTCTG 1604

Qy      65 TTGTCACTTTCCGAGGA 81
      || ||| ||| | |||
Db      1605 TTCTCAGTTTTTCCTGGA 1621
```

RESULT 24

US-10-161-539-3

```
; Sequence 3, Application US/10161539
; Publication No. US20030147854A1
; GENERAL INFORMATION:
; APPLICANT: Gregory, R.J., Armentano, D., Couture, L.A., Smith,
;           A.E.
; TITLE OF INVENTION: ADENOVIRUS VECTORS FOR GENE THERAPY
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENZYME CORPORATION
; STREET: 15 PLEASANT STREET CONNECTOR
; CITY: FRAMINGHAM
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 01701-9322
```

```

;      COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: ASCII
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/10/161,539
;      FILING DATE: 20-Feb-2003
;      CLASSIFICATION: <Unknown>
;      PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: US 09/248,026
;      FILING DATE: 10-FEB-1999
;      ATTORNEY/AGENT INFORMATION:
;      NAME: Newland, Bart G.
;      REGISTRATION NUMBER: 31,282
;      REFERENCE/DOCKET NUMBER: IG4-09.11.2-CON3
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (508) 271-3920
;      TELEFAX: (508) 872-5415
;      INFORMATION FOR SEQ ID NO: 3:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 5635 base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: single
;      TOPOLOGY: linear
;      MOLECULE TYPE: cDNA
;      SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-161-539-3

```

```

Query Match          31.6%; Score 32.2; DB 15; Length 5635;
Best Local Similarity 63.6%; Pred. No. 0.27;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
        || | |||| ||   ||| ||||| ||   | | | ||| || | ||| | || |||
Db      2015 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGGGAAGAATTCATTCTG 2074

Qy      65 TTGTCACTTTCCGAGGA 81
        || ||| ||| | |||
Db      2075 TTCTCAGTTTTCCTGGA 2091

```

RESULT 25

US-09-982-315-3

```

; Sequence 3, Application US/09982315
; Publication No. US20030096762A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, Horst
; APPLICANT: Illek, Beate
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CYSTIC FIBROSIS THERAPY
; FILE REFERENCE: 200116.403D1
; CURRENT APPLICATION NUMBER: US/09/982,315
; CURRENT FILING DATE: 2001-10-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 6126

```


; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-982-315-3

Query Match 31.6%; Score 32.2; DB 10; Length 6126;
Best Local Similarity 63.6%; Pred. No. 0.27;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
          || | |||| ||      ||| ||||| ||      | | | ||| || | ||| | || |||
Db      1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTCATTCTG 1604

Qy      65 TTGTCACTTTCCGAGGA 81
          || ||| ||| | |||
Db      1605 TTCTCAGTTTTCCTGGA 1621
```

RESULT 26

US-09-782-378A-24
; Sequence 24, Application US/09782378A
; Patent No. US20020102731A1
; GENERAL INFORMATION:
; APPLICANT: Hearing, Patrick
; APPLICANT: Bahou, Wadie
; APPLICANT: Sandalon, Ziv
; APPLICANT: Gnatenko, Dmitri
; TITLE OF INVENTION: Adenoviral Vectors
; FILE REFERENCE: STONYB-04970
; CURRENT APPLICATION NUMBER: US/09/782,378A
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/237,747
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 6129
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-378A-24

Query Match 31.6%; Score 32.2; DB 9; Length 6129;
Best Local Similarity 63.6%; Pred. No. 0.27;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
          || | |||| ||      ||| ||||| ||      | | | ||| || | ||| | || |||
Db      1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTCATTCTG 1604

Qy      65 TTGTCACTTTCCGAGGA 81
          || ||| ||| | |||
Db      1605 TTCTCAGTTTTCCTGGA 1621
```

RESULT 27

US-09-982-315-1
; Sequence 1, Application US/09982315
; Publication No. US20030096762A1

```
; GENERAL INFORMATION:
; APPLICANT: Fischer, Horst
; APPLICANT: Illek, Beate
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CYSTIC FIBROSIS THERAPY
; FILE REFERENCE: 200116.403D1
; CURRENT APPLICATION NUMBER: US/09/982,315
; CURRENT FILING DATE: 2001-10-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 6129
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-982-315-1
```

```
Query Match          31.6%; Score 32.2; DB 10; Length 6129;
Best Local Similarity 63.6%; Pred. No. 0.27;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
```

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
      || | |||| ||      ||| ||||| || | | | ||| || | ||| | || |||
Db      1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG 1604

Qy      65 TTGTCACTTTCCGAGGA 81
      || ||| ||| | |||
Db      1605 TTCTCAGTTTTCTGGA 1621
```

RESULT 28

US-09-982-315-5

```
; Sequence 5, Application US/09982315
; Publication No. US20030096762A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, Horst
; APPLICANT: Illek, Beate
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CYSTIC FIBROSIS THERAPY
; FILE REFERENCE: 200116.403D1
; CURRENT APPLICATION NUMBER: US/09/982,315
; CURRENT FILING DATE: 2001-10-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 6129
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-982-315-5
```

```
Query Match          31.6%; Score 32.2; DB 10; Length 6129;
Best Local Similarity 63.6%; Pred. No. 0.27;
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
```

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
      || | |||| ||      ||| ||||| || | | | ||| || | ||| | || |||
Db      1545 TATGGGAGAACTGGAGCCTTCAGAGGGTAAATTAAGCACAGTGAAGAATTTTCATTCTG 1604

Qy      65 TTGTCACTTTCCGAGGA 81
      || ||| ||| | |||
```

Db 1605 TTCTCAGTTTTCCTGGA 1621

RESULT 29

US-10-161-539-1

; Sequence 1, Application US/10161539

; Publication No. US20030147854A1

; GENERAL INFORMATION:

; APPLICANT: Gregory, R.J., Armentano, D., Couture, L.A., Smith,
; A.E.

; TITLE OF INVENTION: ADENOVIRUS VECTORS FOR GENE THERAPY

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENZYME CORPORATION

; STREET: 15 PLEASANT STREET CONNECTOR

; CITY: FRAMINGHAM

; STATE: MASSACHUSETTS

; COUNTRY: USA

; ZIP: 01701-9322

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/161,539

; FILING DATE: 20-Feb-2003

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 09/248,026

; FILING DATE: 10-FEB-1999

; ATTORNEY/AGENT INFORMATION:

; NAME: Newland, Bart G.

; REGISTRATION NUMBER: 31,282

; REFERENCE/DOCKET NUMBER: IG4-09.11.2-CON3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (508) 271-3920

; TELEFAX: (508) 872-5415

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 6129 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 133..4572

; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-10-161-539-1

Query Match 31.6%; Score 32.2; DB 15; Length 6129;

Best Local Similarity 63.6%; Pred. No. 0.27;

Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64

|| | |||| || ||| ||||| || | | | ||| || | ||| | || |||

; PRIOR APPLICATION NUMBER: US 60/218,006

```
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 172129
; LENGTH: 799
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-172129
```

```
Query Match          27.8%; Score 28.4; DB 13; Length 799;
Best Local Similarity 62.9%; Pred. No. 3.9;
Matches 44; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
```

```
Qy      5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
        ||||| | ||| || ||||| || | || | || || || | || || ||
Db      591 TAGGTAATATCAGTGTGCTCCAAAGGTTGAGAATAACTGCTTTAAGTTGAAAAAAGAATG 650

Qy      65 TTGTCACTTT 74
        ||| ||| |
Db      651 TTGGAAGTCT 660
```

RESULT 32

```
US-10-027-632-172129
; Sequence 172129, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
```

; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 172129
; LENGTH: 799
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-172129

Query Match 27.8%; Score 28.4; DB 16; Length 799;
Best Local Similarity 62.9%; Pred. No. 3.9;
Matches 44; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 5 TAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAGACTG 64
||||| | ||| || ||||| || | || | || || || | | ||| ||
Db 591 TAGGTAATATCAGTGTGCTCCAAAGGTTGAGAATAACTGCTTTAAGTTGAAAAAAGAATG 650

Qy 65 TTGTCACTTT 74
||| ||| |
Db 651 TTGGAAGTCT 660

RESULT 33

US-09-925-299-368/c
; Sequence 368, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 368
; LENGTH: 548
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (370)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (378)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (384)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (412)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature

```

; LOCATION: (429)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (449)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (471)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (490)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (495)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (528)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-368

```

```

Query Match          27.6%; Score 28.2; DB 9; Length 548;
Best Local Similarity 61.6%; Pred. No. 4.2;
Matches 45; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy      29 GTGTTGGACTGACCACTGTAGGTGAAGTACAGACTGTTGTCACTTTCCGAGGAGAACAAG 88
          ||||| |||| | |||| || | |||| || ||| ||||| |
Db      327 GTGTTGGTGTGACTATTGTAGCTGGGACATTTACTGTGGTGGGTTTCTGAGGAGTTGGTG 268

Qy      89 CTGTCCTGGAGGC 101
          || || | ||
Db      267 GGGTTCTTGTAGC 255

```

RESULT 34

```

US-09-925-299-368/c
; Sequence 368, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 368
; LENGTH: 548
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (370)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature

```

```

; LOCATION: (378)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (384)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (412)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (429)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (449)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (471)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (490)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (495)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (528)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-368

```

```

Query Match          27.6%; Score 28.2; DB 10; Length 548;
Best Local Similarity 61.6%; Pred. No. 4.2;
Matches 45; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```

Qy      29 GTGTTGGACTGACCACTGTAGGTGAAGTACAGACTGTTGTCACCTTCCGAGGAGAACAAG 88
        ||||| ||| | |||| | | | |||| | | ||| ||||| |
Db      327 GTGTTGGTGTGACTATTGTAGCTGGGACATTTACTGTGGTGGGTTTCTGAGGAGTTGGTG 268

Qy      89 CTGTCCTGGAGGC 101
        || || | ||
Db      267 GGGTTCTTGTAGC 255

```

RESULT 35

US-09-934-814-4/c

; Sequence 4, Application US/09934814

; Patent No. US20020137159A1

; GENERAL INFORMATION:

; APPLICANT: Lok, Si

; APPLICANT: Holloway, James L.

; APPLICANT: O'Hara, Patrick J.

; TITLE OF INVENTION: Human Phermone Polypeptides

; FILE REFERENCE: 00-80

; CURRENT APPLICATION NUMBER: US/09/934,814

; CURRENT FILING DATE: 2001-08-22

; NUMBER OF SEQ ID NOS: 13

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 4

; LENGTH: 432

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(429)
US-09-934-814-4

Query Match 27.5%; Score 28; DB 9; Length 432;
Best Local Similarity 71.2%; Pred. No. 4.7;
Matches 37; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 46 GTAGGTGAAGTACAGACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGG 97
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 366 GAAGGTGATGAACAGCCTGTAGTCAGTCTCCGAGACGGCCACTGTGTTCTGG 315

RESULT 36

US-10-142-465-4/c

; Sequence 4, Application US/10142465
; Publication No. US20030166070A1
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Holloway, James L.
; APPLICANT: O'Hara, Patrick J.
; TITLE OF INVENTION: Human Phermone Polypeptides
; FILE REFERENCE: 00-80
; CURRENT APPLICATION NUMBER: US/10/142,465
; CURRENT FILING DATE: 2002-05-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(429)
US-10-142-465-4

Query Match 27.5%; Score 28; DB 15; Length 432;
Best Local Similarity 71.2%; Pred. No. 4.7;
Matches 37; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 46 GTAGGTGAAGTACAGACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGG 97
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 366 GAAGGTGATGAACAGCCTGTAGTCAGTCTCCGAGACGGCCACTGTGTTCTGG 315

RESULT 37

US-09-934-814-1/c

; Sequence 1, Application US/09934814
; Patent No. US20020137159A1
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Holloway, James L.
; APPLICANT: O'Hara, Patrick J.
; TITLE OF INVENTION: Human Phermone Polypeptides

; FILE REFERENCE: 00-80
; CURRENT APPLICATION NUMBER: US/09/934,814
; CURRENT FILING DATE: 2001-08-22
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(525)
US-09-934-814-1

Query Match 27.5%; Score 28; DB 9; Length 525;
Best Local Similarity 71.2%; Pred. No. 4.9;
Matches 37; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 46 GTAGGTGAAGTACAGACTGTTGTCACCTTTCCGAGGAGAACAAGCTGTCCTGG 97
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 375 GAAGGTGATGAACAGCCTGTAGTCAGTCTCCGAGACGGCCACTGTGTTCTGG 324

RESULT 38
US-10-142-465-1/c
; Sequence 1, Application US/10142465
; Publication No. US20030166070A1
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Holloway, James L.
; APPLICANT: O'Hara, Patrick J.
; TITLE OF INVENTION: Human Phermone Polypeptides
; FILE REFERENCE: 00-80
; CURRENT APPLICATION NUMBER: US/10/142,465
; CURRENT FILING DATE: 2002-05-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(525)
US-10-142-465-1

Query Match 27.5%; Score 28; DB 15; Length 525;
Best Local Similarity 71.2%; Pred. No. 4.9;
Matches 37; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 46 GTAGGTGAAGTACAGACTGTTGTCACCTTTCCGAGGAGAACAAGCTGTCCTGG 97
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 375 GAAGGTGATGAACAGCCTGTAGTCAGTCTCCGAGACGGCCACTGTGTTCTGG 324

RESULT 39
US-09-934-814-7/c

```
; Sequence 7, Application US/09934814
; Patent No. US20020137159A1
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Holloway, James L.
; APPLICANT: O'Hara, Patrick J.
; TITLE OF INVENTION: Human Phermone Polypeptides
; FILE REFERENCE: 00-80
; CURRENT APPLICATION NUMBER: US/09/934,814
; CURRENT FILING DATE: 2001-08-22
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 540
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(537)
US-09-934-814-7
```

```
Query Match          27.5%; Score 28; DB 9; Length 540;
Best Local Similarity 71.2%; Pred. No. 5;
Matches 37; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
```

```
Qy      46 GTAGGTGAAGTACAGACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGG 97
        | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      366 GAAGGTGATGAACAGCCTGTAGTCAGTCTCCGAGACGGCCACTGTGTTCTGG 315
```

```
RESULT 40
US-10-142-465-7/c
; Sequence 7, Application US/10142465
; Publication No. US20030166070A1
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Holloway, James L.
; APPLICANT: O'Hara, Patrick J.
; TITLE OF INVENTION: Human Phermone Polypeptides
; FILE REFERENCE: 00-80
; CURRENT APPLICATION NUMBER: US/10/142,465
; CURRENT FILING DATE: 2002-05-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 540
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(537)
US-10-142-465-7
```

```
Query Match          27.5%; Score 28; DB 15; Length 540;
Best Local Similarity 71.2%; Pred. No. 5;
Matches 37; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
```

Qy 46 GTAGGTGAAGTACAGACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGG 97
 | | | | | | | | | | | | | | | | | | | | | |
 Db 366 GAAGGTGATGAACAGCCTGTAGTCAGTCTCCGAGACGGCCACTGTGTTCTGG 315

RESULT 41

US-09-934-814-10/c
 ; Sequence 10, Application US/09934814
 ; Patent No. US20020137159A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lok, Si
 ; APPLICANT: Holloway, James L.
 ; APPLICANT: O'Hara, Patrick J.
 ; TITLE OF INVENTION: Human Phermone Polypeptides
 ; FILE REFERENCE: 00-80
 ; CURRENT APPLICATION NUMBER: US/09/934,814
 ; CURRENT FILING DATE: 2001-08-22
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 10
 ; LENGTH: 795
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(792)
 US-09-934-814-10

Query Match 27.5%; Score 28; DB 9; Length 795;
 Best Local Similarity 71.2%; Pred. No. 5.5;
 Matches 37; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 46 GTAGGTGAAGTACAGACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGG 97
 | | | | | | | | | | | | | | | | | | | | | |
 Db 366 GAAGGTGATGAACAGCCTGTAGTCAGTCTCCGAGACGGCCACTGTGTTCTGG 315

RESULT 42

US-10-142-465-10/c
 ; Sequence 10, Application US/10142465
 ; Publication No. US20030166070A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lok, Si
 ; APPLICANT: Holloway, James L.
 ; APPLICANT: O'Hara, Patrick J.
 ; TITLE OF INVENTION: Human Phermone Polypeptides
 ; FILE REFERENCE: 00-80
 ; CURRENT APPLICATION NUMBER: US/10/142,465
 ; CURRENT FILING DATE: 2002-05-09
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 10
 ; LENGTH: 795
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS

; LOCATION: (1)...(792)
US-10-142-465-10

Query Match 27.5%; Score 28; DB 15; Length 795;
Best Local Similarity 71.2%; Pred. No. 5.5;
Matches 37; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 46 GTAGGTGAAGTACAGACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGG 97
| | | | | | | | | | | | | | | | | | | | | |
Db 366 GAAGGTGATGAACAGCCTGTAGTCAGTCTCCGAGACGGCCACTGTGTTCTGG 315

RESULT 43

US-10-399-456-3/c

; Sequence 3, Application US/10399456
; Publication No. US20040043395A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: LAL, Preeti G.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LU, Yan
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: BRUNS, Christopher M.
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: YAO, Monique G.
; TITLE OF INVENTION: LIPOCALINS
; FILE REFERENCE: PF-0822 USN
; CURRENT APPLICATION NUMBER: US/10/399,456
; CURRENT FILING DATE: 2003-04-14
; PRIOR APPLICATION NUMBER: PCT/US01/31942
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,541
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 1630
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040043395A1 3537562CB1
US-10-399-456-3

Query Match 27.5%; Score 28; DB 13; Length 1630;
Best Local Similarity 71.2%; Pred. No. 6.6;
Matches 37; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 46 GTAGGTGAAGTACAGACTGTTGTCACTTTCCGAGGAGAACAAGCTGTCCTGG 97
| | | | | | | | | | | | | | | | | | | | | |
Db 366 GAAGGTGATGAACAGCCTGTAGTCAGTCTCCGAGACGGCCACTGTGTTCTGG 315

RESULT 44

US-10-052-482-166

```
; Sequence 166, Application US/10052482
; Publication No. US20040072264A1
; GENERAL INFORMATION:
; APPLICANT: Engelhard, Eric
; APPLICANT: Morris, David
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71087/RMS/DCF
; CURRENT APPLICATION NUMBER: US/10/052,482
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 166
; LENGTH: 48244
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36673)..(36711)
; OTHER INFORMATION: "n" at positions 36673 to 36711 can be any base
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40035)..(40119)
; OTHER INFORMATION: "n" at positions 40035 to 40119 can be any base
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (42958)..(43306)
; OTHER INFORMATION: "n" at positions 42958 to 43306 can be any base
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47841)..(47909)
; OTHER INFORMATION: "n" at positions 47841 to 47909 can be any base
US-10-052-482-166
```

```
Query Match          27.5%; Score 28; DB 12; Length 48244;
Best Local Similarity 58.3%; Pred. No. 16;
Matches 49; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```
Qy          1 CTGGTAGGTGAGATCTCTGACCTCCAGAGTGTGGACTGACCACTGTAGGTGAAGTACAG 60
      ||||  |||  ||  |||  |||||  ||  ||  ||  ||||  |  |||  |  |
Db    40187 CTGGCTTGTGTGTTCTGCCCTCCACTGGGTGCTACGGACCAAGGGCTGTGCTGAGCCC 40246

Qy          61 ACTGTTGTCACTTTCCGAGGAGAA 84
      ||||  ||  ||  ||  ||  ||  |||
Db    40247 CCTGTGGCCGCTCTCACAGCTGAA 40270
```

RESULT 45

US-10-027-632-265948

```
; Sequence 265948, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```



```
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 265950
;   LENGTH: 987
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (1)...(987)
;   OTHER INFORMATION: n = A,T,C or G
US-10-027-632-265950
```

RESULT 48

US-10-027-632-265948

Query Match 27.3%; Score 27.8; DB 16; Length 987;
Best Local Similarity 62.0%; Pred. No. 6.9;
Matches 44; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

```
Qy      29 GTGTTGGACTGACCACTGTAGGTGAAGTACAGACTGTTGTCACTTTCCGAGGAGAACAAG 88
          ||| |||  || || || ||  || | ||||| |  |||| ||  ||||| |  ||
Db      253 GTGGTGGTAGGAGCATTGCAGAGGATGGACAGATTTCATGTCCCTCAGAGAGGAGGAGGAG 312

Qy      89 CTGTCCTGGAG 99
          ||| |||
Db      313 AAGGCATAGAG 323
```

RESULT 50

US-10-027-632-265949

; Sequence 265949, Application US/10027632

; Publication No. US20030204075A9

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; TITLE OF INVENTION: Polymorphisms in the Human Genome

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 265949

; LENGTH: 987

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(987)

; OTHER INFORMATION: n = A,T,C or G

US-10-027-632-265949

Query Match 27.3%; Score 27.8; DB 16; Length 987;
Best Local Similarity 62.0%; Pred. No. 6.9;
Matches 44; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

```
Qy      29 GTGTTGGACTGACCACTGTAGGTGAAGTACAGACTGTTGTCACTTTCCGAGGAGAACAAG 88
```

```

      ||| |||  || || || ||  || | |||| |  |||| ||  ||||| |  ||
Db      253 GTGGTGGTAGGAGCATTGCAGAGGATGGACAGATTCATGTCCCTCAGAGAGGAGGAGGAG 312

Qy      89 CTGTCCTGGAG 99
      | | | |||
Db      313 AAGGCATAGAG 323

```

Search completed: April 29, 2004, 21:08:43
Job time : 100.194 secs